AD NUMBER

ADB219964

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Proprietary Info.; Oct 96. Other requests shall be referred to Commander, Army Medical Research and Materiel Command, Attn: MCMR-RMI-S, Fort Detrick, Frederick, MD 21702-5012.

AUTHORITY

USAMRMC ltr, 1 Jun 2001.

AD_____

GRANT NUMBER DAMD17-94-J-4391

TITLE: Isolation of a Breast Cancer Tumor Suppressor Gene From Chromosome 3p

PRINCIPAL INVESTIGATOR: Harry A. Drabkin, M.D.

CONTRACTING ORGANIZATION: University of Colorado, Denver
Denver, Colorado 80262

REPORT DATE: October 1996

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1996	Annual (23 Sep 95 - 22 Sep 96)

4. TITLE AND SUBTITLE Isolation of a Breast Cancer Tumor Suppressor Gene From Chromosome 3p

5. FUNDING NUMBERS
DAMD17-94-J-4391

6. AUTHOR(S)
Harry A. Drabkin, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Colorado, Denver
Denver, Colorado 80262

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick
Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution authorized to U.S. Government agencies only; Proprietary Information, Oct 96. Other requests for this document shall be referred to Comander, U.S. Army Medical Research and Materiel Command, ATTN: MCMR-RMI-S, Fort Detrick, Frederick, MD 21702-5012

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

Loss of tumor suppressor genes (TSGs) represent critical molecular events in the development and progression of breast cancer. Based on loss of heterozygosity (LOH) studies as well as direct cytogenetic studies of breast tumors, one or more TSGs likely resides on the short arm of chromosome 3 (3p) and appears to be involved in nearly 50% of breast cancers. Four distinct regions within 3p [p12, p14, p21 (proximal) and p21 (distal)] undergo recurrent deletions in human carcinomas and are the most likely sites for a breast cancer TSG. Recently, we demonstrated recurrent homozygous deletion or rearrangement in breast cancer cell lines involving 3p14. A set of DNA clones spanning the critical region has now been sequenced revealing several potential gene coding segments. YACs containing the critical region have been modified with the Neomycin resistance gene for selective retention in mammalian cells. Introduction of YACs into a homozygously deleted breast cancer cell line should provide direct evidence concerning the presence of a functioning tumor suppressor gene. The progress made in 3p14 has now allowed us to evaluate the other candidate gene regions for involvement in breast cancer as originally proposed using deletion detection, DNA sequence analysis, gene isolation and expression assays.

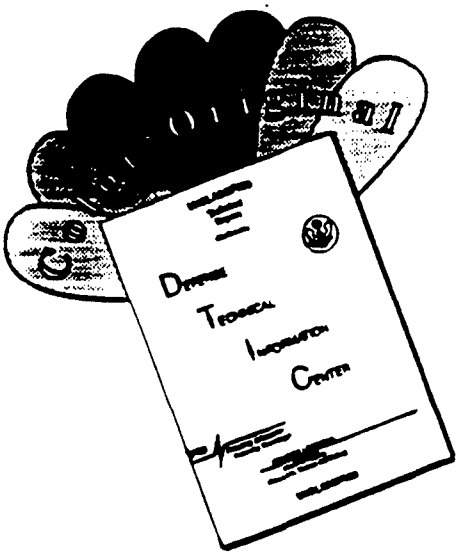
14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	65
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Limited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF COLOR PAGES WHICH DO NOT REPRODUCE LEGIBLY ON BLACK AND WHITE MICROFICHE.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

NO Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

NO In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

NO For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

NO In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

NO In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Gary D. Drake 10/16/96

PI - Signature Date

TABLE OF CONTENTS

	Page #
Front Cover	1
SF 298, Report Documentation Page	2
Foreword	3
Table of Contents	4
Introduction	5 - 9
Body *	9 - 16
Conclusions *	16 - 17
References	17 - 20
Appendices *	21 - 65

* = Unpublished Data

5. INTRODUCTION

5.A. Purpose of the Present Work

Our project concerns the identification and isolation of a breast cancer tumor suppressor gene from the short arm of chromosome 3 (3p). In the early stages of this investigation, we identified a region of homozygous deletion in a subset of breast cancer cell lines suggesting that a tumor suppressor gene would be found in the region of loss. The deletion occurred in a segment of DNA within a few hundred kilobases (1 kb = 1,000 base pairs, a measure of distance along the DNA molecule) of a chromosomal rearrangement involving chromosomes 3 and 8 which is associated with hereditary kidney carcinoma (1). (It would not be uncommon for a tumor suppressor gene to be involved in more than one type of cancer, and both kidney and breast cancers are of epithelial cell origin.) The region surrounding this breakpoint (located in band 3p14.2) was one of the possible target loci described in our original application. Our finding of homozygous deletions near this site caused us to temporarily modify the order of experimentation to permit a detailed examination of this region. This allowed us to develop necessary DNA reagents to study the target area thoroughly and to identify potential genes. We have made significant progress in this regard assembling the completed sequence of over 150 kilobases of DNA and identification of a number of potential gene coding regions (exons). We have also submitted a first manuscript on these results (2). While we continue to study the 3p14 region, we have now been able to include efforts directed at the three additional sites which may harbor important breast cancer related genes. The Progress Report first describes our results in the 3p14 homozygous deletion region and then covers new findings in these other three regions.

5.B Nature of the Problem

Basic concepts especially for the lay reader We realize that the scientific literature is nearly totally composed of technical terms. Throughout this report we have attempted to explain these concepts in lay terms.

The malignant potential of any tumor, including breast cancer, is a consequence of specific alterations (mutations, deletions, amplifications, etc.) in target genes that regulate the growth and biologic behavior of those cells. (Genes are segments of DNA which encode proteins; DNA is "transcribed" into RNA and RNA in turn is "translated" into protein.) Whether cells grow slowly and remain localized, or proliferate rapidly and spread to distant sites (metastasize) is a complex process involving a host of regulatory genes. For example, loss or mutation of the p53 tumor suppressor gene, located on 17p, is associated with instability of the genome (entire DNA of the cell) (3,4) and a worsened prognosis. This instability results in an enhanced capacity of the malignant cell to undergo DNA rearrangements leading to alterations in critical regulatory genes. Loss of normal p53 function is also associated with the cell's ability to escape death or cell cycle arrest resulting from therapeutic radiation or chemotherapy (5).

In other instances, the critical regulatory genes have yet to be identified. This is the case for genes located on 3p although we now have candidates in target regions. Scientific investigations have provided strong evidence pointing to where certain types of critical genes are likely to be located. For example, cytogenetic studies, which examine the content and nature of chromosomes within cells, have identified certain recurrent abnormalities in cancers. Specific chromosomal segments have been found to be increased in number (amplified). This finding is expected to be associated with overexpression of a gene (because of its increased copy number). Such genes, for example, may encode growth factor receptors or may encode proteins that mediate resistance to chemotherapeutic agents. An example is provided by the MDM2 gene whose protein product inhibits the activity of p53; overexpression of MDM2 is thought to have consequences similar to mutation in p53. In contrast, cytogenetic studies have also pointed to recurrent deletions involving specific chromosomal regions. The critical genes believed to be encoded in these regions are

referred to as tumor suppressor genes, the type of gene located on 3p which is the focus of our investigation.

The nature of known tumor suppressor genes is quite varied. Certain tumor suppressor genes, e.g. p16 (an inhibitor of the cyclin dependent kinases or CDKs) and the retinoblastoma gene (RB1, an inhibitor of the E2F transcription factor), control cell division by regulating the process of DNA replication. Some tumor suppressor genes, such as the chromosome 18 gene DCC (Deleted in Colon Carcinoma), are located on the cell surface. Excitingly, as this may relate to our studies on the Semaphorin IV gene located in the proximal 3p21.31 deletion region, DCC has recently been shown to encode a receptor for a netrin (unpublished results), a protein involved in nerve growth cone development. This demonstrates that such molecules initially identified in signaling pathways associated with nerve growth cone guidance can be involved in cancer as we discuss later. A common feature in this class of tumor suppressor genes is that their normal function is lost as part of tumor development. As a consequence, a regulatory function generally affecting growth and differentiation is also lost. The resulting cell may divide more frequently than is appropriate, giving rise to a clone or small cluster of related but abnormal cells and setting the stage for further genetic changes.

5.C. Methods to Isolate/Identify Tumor Suppressor Genes

5.C.1 Positional Cloning.

Tumor suppressor genes have been isolated by two approaches. Perhaps the most frequently used method is referred to as "positional cloning" in which the region of chromosomal loss is defined by molecular (DNA) probes and cytogenetic analysis. Because a visible chromosome deletion represents a large expanse of actual DNA, it is necessary to narrow the target region as much as possible. In one approach, this is done by using "polymorphic" probes which can distinguish between the two copies of the chromosome in question (each being inherited from one parent). In the tumor DNA, loss of one copy (by a variety of mechanisms) is referred to as "loss of heterozygosity" (abbreviated LOH). To detect LOH, a DNA probe must exist which corresponds to the target DNA in question and, importantly, this bit of DNA must exhibit a frequent, naturally occurring, variation in the population. Such variations can usually be detected experimentally and they provide a means to identify differences between the two chromosome copies in any individual tumor sample. Naturally occurring differences in DNA sequence are not uncommon although some types of DNA sequences, referred to as "microsatellites", exhibit much more variation than other types. Microsatellites consist of repeated pairs of nucleotides (usually cytosine followed by adenosine, abbreviated CA) at specific chromosomal sites with the important feature that the number of pairs is variable from person to person. Since the natural variation of markers is often a limiting factor even with microsatellite markers, this explains the general necessity to test fairly large numbers of tumor samples with different DNA probes. The goals of LOH experiments are to identify one or more target regions and to narrow such regions as much as possible prior to performing gene searches.

5.C.2 Candidate Genes.

A second approach in identifying tumor suppressor genes is through the testing of candidate genes. These candidates come from two sources; first, genes mapped within regions of recurrent deletion (which is essentially the positional cloning approach except that the genes have already been identified) are all considered candidates until proven otherwise, and second, those candidates whose known function suggests they might be targets independent of chromosomal position. While our studies continue to involve the positional cloning approach, a number of candidate loci have now been identified in the target regions and we are beginning to test these. These genes include several putative coding regions within 3p14.2, the semaphorin IV gene in 3p21.31 and at least 5 cDNAs now identified in 3p21.33.

5.C.3 Identification of Homozygous Deletions Greatly Facilitates Positional Cloning of Tumor Suppress Genes

LOH studies can lead to the identification of homozygous deletions in tumors, an extremely important finding which can greatly facilitate the precise positioning of tumor suppressor genes. Furthermore, homozygous deletions provide one of the best ways to chose particular candidate genes from a larger set, based upon their location within the minimal deletion defined experimentally. These often rare events thus provide a powerful adjunct to LOH and candidate gene studies.

A homozygous deletion means that both copies of a chromosome have undergone loss for a particular segment of DNA (see diagram below).

Homozygous Deletion in Tumor

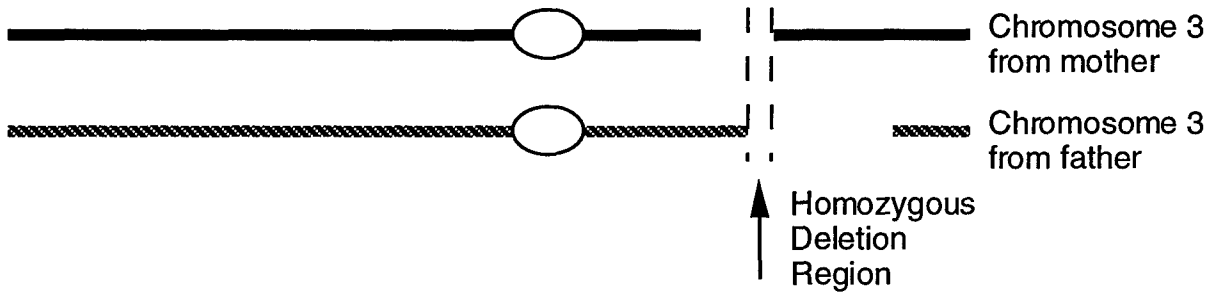


Figure 1. Diagram of chromosome 3 homologues with overlapping deletions in a tumor.

Such deletions represent one of several mechanisms that can lead to the complete loss of a tumor suppressor gene. While tumors frequently lose one tumor suppressor gene copy by undergoing large chromosomal deletions, the remaining copy is usually rendered non-functional by other genetic mechanisms. These can include point mutations and gene silencing (caused by hypermethylation) as well as a second deletion leaving a complete gap in the genetic material.

5.C. Additional Background Information.

In this section, we present a summary of data from the literature regarding genetic deletions of chromosome 3 in breast cancer. These selected studies relate both to our identification of a homozygous deletion in 3p14 as well as genetic changes in other regions. We have placed these comments here for the convenience of the reader, since the concepts of loss of heterozygosity (LOH) and homozygous deletions have been introduced earlier.

Sato et al. (6) examined 120 breast cancers for loss of heterozygosity (LOH) using a series of 3p polymorphic loci. Nearly 50% of the informative tumors (56/120) demonstrated LOH involving 3p. In this study, the 3p region undergoing the greatest loss was 3p14.2-p13 which should include the homozygous deletion region we have identified. This study has been one of the largest in terms of the number of tumors examined and 3p-derived probes tested. One significant limitation was that the type of probe used did not detect sufficient polymorphisms to allow the limits of the 3p target region to be defined more precisely. Another study using LOH analysis demonstrated that 3p loss was most frequently seen in familial breast cancers (7). Excitingly, the highest frequency of loss (68%) was with a probe (D3S1217) also located in the 3p14 region. A cytogenetic study (chromosome analysis by direct microscopic visualization) by Pandis et al. (8) identified 3p deletions in 5 of 41 breast carcinomas. Importantly, the deletions appeared nearly identical and involved the 3p13(p14) region implicated above using LOH analysis. Intriguingly, in

the study of Pandis et al., in 3 of the 5 cases the 3p deletion was seen as the only recognizable cytogenetic abnormality. This suggests that 3p deletion may be an early event in a subgroup of breast cancers. Taken together, there is considerable evidence from multiple investigators using different techniques which supports the frequent alteration of 3p14 during breast cancer development.

Additional data from LOH and cytogenetics studies strongly implicate several other 3p regions that may harbor genes important in breast cancer. Chen et al (9) found LOH within 3p13-14, 3p21-22 and 3p24-26 along with one case of homozygous deletion in 3p13 likely corresponding to our homozygous deletion. All three of these target sites have been previously implicated in other carcinomas, particularly of lung. Hainsworth et al., (10) found frequent chromosomal breakpoints at 3p21 which can be indicative of gene disruption or deletion. In a study by Ali et al, (11), the shortest region of common loss was between the markers D3F15S2 and RAF1, spanning the chromosomal region 3p21-p25 and a similar study by Devilee et al (12) found the p14 to p21 region commonly deleted. Several other studies corroborate these findings having observed interstitial 3p deletions (13) or other structural changes affecting this chromosome (14). Intriguingly, Buchagen et al., (15) identified one apparent rearrangement and one homozygous deletion for the probe D3S2, known to map within proximal 3p21.1 (16) very close to the semaphorin IV deletion region (17). These studies all implicate multiple target regions on 3p which appear to be very similar to regions identified from studies on carcinomas derived from other tissues. Thus as our investigation of 3p14 has matured, we are pursuing these additional regions with the markers and gene sequences available as well as isolating new materials for this purpose.

5.D. Review of Specific Aims

5.D.1. Original Specific Aims.

Briefly, our original Specific Aims were designed to:

1. Define the regions of 3p undergoing LOH in breast cancer.
2. Test known 3p candidate genes for mutations in breast tumors.
3. Isolate additional candidate tumor suppressor genes from regions we define.
4. Characterize the product(s) of these genes and assess their involvement in breast cancer.

5.D.2. A Modified Approach

The above aims were written before we discovered a recurrent homozygous deletion involving 3p14 in a subset of breast carcinoma cell lines. In the last Progress Report, we modified the aims to reflect the substantial effort being focused on this key region. Gene identification (our third original aim) in this region became our highest priority along with testing for deletions, mutations and expression in other breast cancer cell lines and direct tumors. Secondarily, we expected to develop a number of markers that could be used for surveys of additional tumors and cell lines to improve our knowledge of target regions (our first original aim).

5.D.3. Current Strategy.

We have now achieved many of the goals set out in last year's progress report for the 3p14 region. Detailed results from this investigation are presented below and have been submitted for publication (2). This progress has allowed us to pursue three other sites where potential breast cancer tumor suppressor genes may be found, following the original strategy and aims of our proposal. We are testing breast cancer lines for deletions in these regions and are collaborating with investigators in Iceland to test tumors. A large number of non-polymorphic markers have been developed in these regions which are being used for a saturation screen looking for homozygous

deletions. In addition, we have identified a number of polymorphic microsatellite probes (which allow frequent discrimination between the two parental copies of chromosome 3) from each deletion target region. These will be used to ascertain LOH in direct tumors (in contrast to cell lines). The DNA sequence analysis necessary for marker development has identified a number of potential genes which are being tested for alterations in breast cancers. We have also developed an antibody against the semaphorin IV protein encoded within the 3p21.31 deletion region and tested breast cancers for expression. The ultimate test of tumor suppressor gene function is to replace the gene in question in appropriate tumor cells and determine what effect the expressed gene has on growth and differentiation. We are doing this for the 3p14 deletion and will begin such studies for the other regions in the next period of support.

6. BODY (UNPUBLISHED DATA)

6.A. Homozygous deletions at 3p14 in breast carcinomas.

6.A.1. Examination of breast cancer cell line DNAs with probes from 3p14.

From LOH analysis in a variety of carcinomas, one of the most frequently lost regions is 3p14 (18). As already mentioned, this region is also of interest since it contains the site of a hereditary kidney carcinoma associated chromosomal translocation, t(3;8)(p14.2;q24.1) (1), and is the location of the most inducible common fragile site in the genome, FRA3B (19). Fragile sites are regions in the genome that are unstable. In a few cases of rare fragile sites, their nature has been elucidated at the DNA sequence level and appears to be due, at least in part, to an expanded triplet repeat which may interfere with normal DNA replication. As discussed in our previous Progress Report, Southern blot hybridizations were used to detect homozygous deletions and or rearrangements with probes from 3p14 and DNAs from breast cancer cell lines. The breast carcinoma cell lines (n=13) utilized to date are: CRL1504, HTB122, HTB23, HTB121, HTB123, HTB127, HTB131, HTB132, MCF7, MDA231, T47DV, ZR751 and HTB126. However, other tumor cell lines have also been examined which have provided critical information for this deletion.

We have expanded our analysis of this region by developing a ~300 kb cosmid/lambda (DNA clone) contig in the region of the homozygous deletions (Fig. 1). These clones provided us with more probes to detect deletions and provided the source for DNA sequencing studies. From our data, as well as various aphidicolin induced breakpoints and plasmid or viral integration sites (Fig. 1), FRA3B represents a region rather than a single site. (Aphidicolin is a DNA polymerase inhibitor and especially induces FRA3B.) Where we have accurately defined the boundaries for the carcinoma-associated deletions, one or both boundaries are contained within the FRA3B region. Thus, the deletions overlap FRA3B. We have also determined the complete DNA sequence for approximately 150 kb within the region for two purposes: to identify genes and as a means to understand the instability of this region.

During our studies, Ohta et al. (19) identified the FHIT gene with reported abnormalities in RT-PCR products involving a variety of carcinoma cell lines, especially colorectal and lung carcinomas. However, these investigators have demonstrated abnormalities in breast cancer cell lines as well. (In RT-PCR experiments, RNA is isolated and reverse transcribed into cDNA. PCR primers are then used to greatly amplify this product using the polymerase chain reaction which can then be examined for its correct size or its DNA sequence can be determined.) While FHIT could be the putative tumor suppressor gene, we think that it is not for many reasons. First, the smallest deletion we identified (in cell line CC19) does not involve FHIT coding sequences (Figs. 1 and 2A) and from our RT-PCR and cDNA sequence analysis the coding portion of the FHIT transcript is normal in this cell line. This excludes the possibility of an undetected mutation in the coding portion of FHIT. Second, we have observed that FHIT undergoes alternative splicing in normal tissues (Fig. 2, part B) which explains some previously reported abnormal PCR products (20-22). This hypothesis is supported by the recent report of Thiagalingam et al. (23) who observed the

lack of FHIT involvement in colorectal carcinomas and suggested that PCR artifacts might be responsible for some observed alterations (20,21). However, we note that Thiagalingam et al. (23) would have missed many deletions since only few markers were tested. Similarly, their work offered no explanation as to the nature of the deletions and also did not seek to identify alternate candidate genes. Third, discontinuous deletions appear common in this region, both from our analysis and from that reported by Ohta et al. (20). We believe that discontinuous deletions would be observed if they occurred in a region of underlying instability instead of being biologically selected for. Similarly, we have observed deletions in DNAs from non-tumor sources (again suggesting the absence of a biologic selection). Fourth, FHIT has similarity to yeast di-adenosine hydrolase (a metabolic enzyme) which would represent an unexpected function for a tumor suppressor gene. Fifth, and very importantly, re-introduction of FHIT into cell lines where it was altered has had very little consequence (Dr. Carlo Croce, personal communication). Sixth, one of the possible features that suggested FHIT could be a tumor suppressor gene was that it crossed the hereditary kidney carcinoma associated 3;8 breakpoint (20). However, we found no alterations in RT-PCR products from 5 renal carcinoma cell lines, and Dr. Gyula Kovacs (Heidelberg, Germany) has informed us that they have observed normal FHIT transcripts with no point mutations in a large series of kidney cancers.

Thus, we are faced with two contrasting possibilities. While another candidate tumor suppressor gene may exist within FRA3B, and this needs to be carefully explored as we are doing, an alternative possibility is that the deletions are due to primary genomic instability affecting a particularly susceptible region. In fact, we have observed the highest frequency of deletions in cervical carcinomas where p53 inactivation is very common (24). In this regard, it is interesting that we have also observed an inverse correlation between 3p14 deletions and microsatellite instability which is due to mutations in DNA mismatch repair genes. It has been reported that p53 mutations, which destabilize the genome (3,4), are uncommon in colorectal carcinomas and gastric tumors with microsatellite instability ((25,26) and P. Cottu, presented at Cancer and the Cell Cycle, Lausanne, Switzerland 1996). These findings would be consistent with 3p14 deletions resulting from genomic instability accompanying p53 inactivation. A somewhat recent report described the presence of EB viral sequences in breast cancers. Because EBV might destabilize the host genome, we examined our cell lines by PCR for EBV and found them all negative. Thus, our best candidate for causing genomic instability at the present time is p53 mutation.

6.A.2 DNA Sequencing and Analysis

As stated, we undertook DNA sequencing studies as a means to identify genes and to understand the instability of the region. To date, we have completed the DNA sequencing for approximately 150 kb (delimited in Fig. 1) and our analysis of the sequence features is complete for a contiguous stretch of 110 kb (Fig. 3). (The DNA sequence analysis involves computer algorithms to identify all repetitive DNA sequences, identities and similarities to known genes contained in various databases, predicted gene segments and various other structural/compositional features.) We find the region to be very high in A-T content with frequent LINE and MER repetitive elements, and it is conversely low in Alu repetitive elements and confirmed genes. In contrast to the reported rare folate-sensitive fragile sites, which are associated with expanded CGG repeats (27,28), FRA3B (a common fragile site) does not contain an expanded triplet repeat or methylated CpG-island. Nor did we identify any telomeric repeats which have been suggested as a possible cause of chromosomal breaks (29). However, we did identify a very significant similarity ($1e^{-77}$) to a small polydispersed circular (spc) DNA (position 106 kb, Fig. 3). This spcDNA, which appears non-repetitive, was isolated from a tuberous sclerosis-associated angiofibroma (Ph.D. thesis, I. Hinkel-Schreiner). Characteristics of spcDNAs include derivation from chromosomal sequences (30), association with clustered repeats (such as β-satellites and other clustered elements although a single family member may be predominantly involved (31)) and elevation in conditions associated with genomic instability such as Fanconi's anemia (32), a syndrome which has a marked predisposition to develop cancer. spcDNAs are also increased by

DNA damaging agents (30) and inhibitors of DNA and protein synthesis including the fragile site inducer aphidicolin (33). Given the limited number of spcDNAs that have been sequenced, this similarity may be biologically important as relates to the instability of the region. While we specifically do not know if the spcDNA site is a primary cause of FRA3B instability or simply a marker for this property, the DNA sequence data should provide a means to test this.

With respect to other possible target genes, from our sequencing studies we identified a 100% identity to two expressed-sequence tagged sites (ESTs) from a liver/spleen cDNA library. (ESTs are short DNA sequences obtained from the ends of cDNA clones. While most cDNA clones represent parts of genes, there are some which are derived from incompletely spliced RNA and others which may represent DNA contamination in the library.) However, we determined that both clones were identical, were not expressed using a commercial Northern blot (Clontech), were co-linear with genomic DNA containing a polyA tract corresponding to their 3' end, encoded no significant open reading frame and overlapped a partial LINE element. Thus, this EST does not correspond to a real gene as best we can tell.

Although our analysis of the sequence data and experimental testing is not yet complete for the segment encompassed by cosmid 139C6 (Fig. 1), its sequence appears less AT-rich and contains exon 5 of the FHIT gene. We have also identified a marked similarity (or identity if minor differences are due to DNA sequencing errors) to an EST in this region and have obtained the corresponding cDNA clone from Research Genetics to determine if this represents a real gene, pseudogene or related gene. Thus, it is possible that more genes than FHIT are encoded in the 3p14 homozygous deletion region.

6.A.3 Introduction of YACs into recipient cells.

As a functional assay and as an alternative means to identify potential tumor suppressor genes within large cloned DNA segments, we proposed to introduce YACs into recipient cell lines and then to assay the tumorigenicity of the resulting cells in immunodeficient (nude) mice. This represents specific aim 4 as applied to this chromosomal segment. In our previous report, we described progress in the introduction of a selectable marker (retrofitting) into YAC 74B2g with pRV1 (neoR vector) and the verification of the correct integration. This YAC, containing approximately 450 kb of human DNA from 3p14, represents a portion of the deleted segment in the breast cancer cell line MDA231. To test the system initially, we chose mouse A9 cells for several reasons. First, other investigators have more frequently succeeded in introducing YACs into rodent cell lines. Second, we could easily examine recipient cells to determine if the YAC had been transferred intact or if it had been disrupted. Third, mouse A9 cells are tumorigenic in nude mice and we could ask whether any genes contained in YAC 74B2g affected this.

We have succeeded in transferring the intact YAC to A9 cells. To accomplish this, the retrofitted YAC was purified by preparative pulsed-field gel electrophoresis along with agarase treatment (to remove agarose), and the intactness of the YAC DNA was re-verified. Two-hundred ul of purified YAC DNA was mixed with 50 ul of "Lipofectin", the mixture added to A9 cells and neomycin (G418, 400 ug/ml) resistant clones were subsequently selected. Altogether, 12 resistant clones were obtained. The intactness of the integrated YAC was assessed first by PCR and subsequently by fluorescent in situ hybridization (FISH) using various DNA probes throughout the YAC. For the initial PCR analysis, primers from the vector portion of the YAC were used. While the neoR marker (right arm of vector) gave positive signals in all of the subclones, only two (clones #10 and #12) were definitely positive for the left arm marker. This indicated that in these two clones the YAC was likely intact. These results have been confirmed by FISH analysis using two cosmid clones (31E1, 55D12) locate towards the ends of the human insert in the YAC (Fig. 4). The resulting A9 clones will now be injected into nude mice to assess the effect of the introduced YAC on tumorigenicity.

Based on the results and experience we gained from the A9 transfection experiments, we have now performed similar experiments with the breast carcinoma cell line MDA231. The Lipofectin

transfection process has been completed and we are waiting for the emergence of G418 resistant clones.

As an alternative and improved means of introducing YACs, Dr. Claire Huxley has demonstrated that a linear YAC can be circularized using a vector which contains the replication origin for the EB virus (oriP) and subsequently transferred to mammalian cells. For cells expressing the virally encoded gene EBNA1, which can be transfected into any recipient cell line, the YAC will replicate autonomously as an episome (extrachromosomal). Because the YAC replicates as an episome and does not integrate into the host cell chromosomes, there appears to be far fewer structural alterations and no interference, due to chromosome position effects, to the expression of genes contained in the YAC. We believe this approach will allow us to introduce far larger YACs, and consequently more genes, into recipient cells. Perhaps the only disadvantage to the system is that multiple and variable number of YAC copies are present in each cell. However, we feel the advantages of this system are significant and Dr. Huxley has agreed to provide us with the necessary vectors which are expected shortly.

6.A.4. Summary and Significance of 3p14 Findings

The identification of a DNA segment which undergoes recurrent homozygous deletion and rearrangement, as we have observed for this region of 3p14, suggests that it encodes a tumor suppressor gene. However, our studies have led us to suggest an alternative hypothesis; that genomic instability per se may lead to the observed deletions. These are contrasting hypotheses which can only be answered by careful study. Clearly, at the present time our data do not support FHIT as a tumor suppressor gene. We have identified possible other genes in the deletion region which are under active exploration. Our studies in this region will continue in the current direction of gene identification and analysis of the deletions and their sequence. We suspect that p53 mutation could be responsible for the observed instability, but this must be verified. To investigate this possibility, we are implementing a rapid yeast-based mutation detection system developed by Dr. Richard Iggo (ISREC, Espalinges Switzerland) in order to correlate the presence or absence of 3p14 deletions/rearrangements with p53 mutation. For a functional assay of tumor suppression, we are introducing YACs from this region back into cell lines that show deletions. Our studies of this region address each of the 4 specific aims. Clearly this is an identified region of loss (aim 1). The analysis of FHIT DNA coding sequences has failed to support its involvement as a tumor suppressor gene (aim 2). Our search for additional genes is very much ongoing (aim 3) and finally, the introduction of YACs from this region (aim 4) should provide independent functional evidence for the presence or absence of a tumor suppressor gene.

6.B. Involvement of Proximal 3p in Breast Carcinoma.

6.B.1 Analysis of the 3p12-13 deletion region.

Polymerase chain reaction (PCR) analyses were used to assess the breast carcinoma cell lines for homozygous deletions in the 3p12 area that corresponds to a homozygous deletion identified in the SCLC cell line U2020 (34,35) and in one fresh tumor (36). 3p12 is part of a larger region implicated as a site for a likely tumor suppressor gene through functional analysis for tumor suppression by Sanchez et al (37) and was one of the target regions in our original proposal. PCR reactions were performed using a touch down procedure in which the primer annealing reaction begins at high temperatures (eg., 65°C) and is progressively lowered by 0.5 degrees to the optimum temperature (eg., 55°C) during the first 20 PCR cycles. The final 15 cycles are then performed at the optimum temperature. This procedure yields highly specific amplification products with a minimum of false background bands. Each reaction was performed using 40 nanograms (ng) of cell line DNA in a 20 µl reaction volume. In addition, markers were amplified concurrently (ie., multiplexed) so that absence of a product could not be attributable to a failed PCR reaction. PCR results were obtained for 12 breast carcinoma cell lines; CRL-1504, HTB-122, HTB-23, HTB-121, HTB-123, HTB-127, HTB-131, HTB-132, MCF-7, MDA-231, T-47-D, and

ZR-75.1, using 7 polymorphic markers from within the 3p12 deletion region. The seven markers utilized were D3S1776, D3S1254, AFM191vd8, D3S1577, D3S1604, D3S1274, AFM001za2, listed in their order along the chromosome from proximal (close to the centromere) to distal (towards the telomere, the structure which caps the chromosome arm). Two of these markers (D3S1776 and D3S1254) have also been found to be homozygously deleted in a small-cell lung tumor sample. Control amplifications included DNA from the normal human cell line FS and a reaction with no template to control for contamination. Examples of this analysis for two markers are shown in Fig. 5 in which the products have been detected after separation by agarose gel electrophoresis. All the cell lines had amplifiable products from each marker used in the PCR amplification reaction. We found no evidence of homozygous deletions with this set of cell lines and markers. However, the density of markers is relatively low in this region (about 1 per megabase). From our analysis of a deletion in a direct SCLC tumor, we have narrowed the target region and are preparing more sequenced tagged sites (STS) to permit a higher density screen.

To extend these results distally, we analyzed the same cell lines with three additional markers (D3S1284, AFM320yb5 and D3S1210) which map just telomeric to the U2020 deletion region. These markers lie adjacent to the 3p13 marker D3S1217 which was found by Bergthorsson et al., (7) to show the highest frequency of allelic imbalance in breast tumors from Icelandic kindreds. Again, no homozygous deletions were detected with this set of markers. However, it will be necessary to carry out further tests of tumors with many more markers from this region before we can eliminate it as a likely site for a gene important in breast cancer. This continuing analysis is important because of the suggestive evidence that an important gene may map within the proximal portion of 3p.

6B.2 Significance

This regions represents a putative target from work in other carcinomas. To date, we have observed no homozygous deletions. However, before excluding the region as a strong candidate we will examine a subsegment in greater detail, which we feel has the highest probability of involvement. These studies relate to specific aim 1.

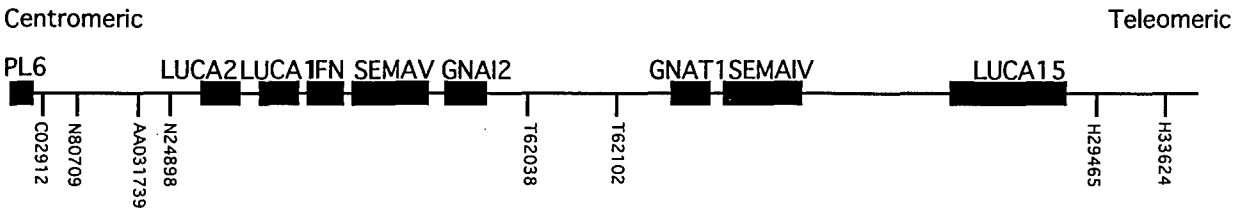
6.C. Involvement of 3p21.31 and the Semaphorin IV Gene in Breast Carcinoma.

6.C.1. Known and Predicted Genes in Region.

Chromosome band 3p21 has long been thought to contain a tumor suppressor gene based upon two seminal observations. First, the region has the highest frequency of LOH in many carcinomas, particularly lung cancer, and has been the target of homozygous deletions in cell lines and uncultured tumors. Importantly, the involvement of this region has been repeatedly suggested by LOH studies in breast cancer. Secondly, functional studies by Killary et al., (38) have shown that DNA segments from this region will suppress tumorigenicity of the mouse fibrosarcoma cell line, A9.

Recently, Lerman et al., (39) isolated a contiguous set of cosmid and P1 clones which span over 600 kb of this region. These clones have been partially sequenced by the Washington University Genome Center and the results have been posted on the Internet. Thirty-two individual sequences totalling ~400 kb from the 3p21.31 region were downloaded and examined. Many of the sequences appeared to be wholly or partially derived from *C. elegans*, *E. coli*, or various vectors, and these sequences were discarded (a total of about 50 kb). The remaining 350 kb was checked for database homologies, common repeat sequences and predicted exons. Repetitive sequences were removed by screening against a repeat sequence database and exons were predicted using GRAIL2 and GeneMark software. BLAST homology searches revealed identities to nine known genes (black boxes) whose relative positions are shown below. Further, a number of

essentially perfect matches to unannotated cDNA sequences were found, many of which coincided with strongly predicted exons.



6.C.2. Analysis by RT-PCR for Gene Expression and Homozygous Deletions in Breast Carcinomas.

A systematic search within the 3p21.31 region for genes whose expression is altered or extinguished in breast tumors is ongoing. We chose to utilize RT-PCR to analyze breast tumor RNAs for gene expression. Primer pairs to date have been designed to amplify all nine known genes and six of the novel gene segments that were highly homologous to cDNAs in the dbEST database. Where possible, primers were chosen within different exons to aid in the discrimination between amplification products originating from genomic DNA vs. messenger RNA. While these results are preliminary and ongoing, we have observed the absence of expression in two closely positioned genes, LuCa 2 and the IFN-related gene in one cell line (Fig.6). However, the RT-PCR results appear problematic with this sample since the positive control (GAPDH) did not work. However, products from 3 other genes were positive (PL6, LuCa 15 and CO2912). Transducin, which is specifically expressed in the retina was silent, as expected, in these samples. We will continue and expand this analysis.

6.C.3. The H.SemaIV gene.

Using a positional cloning strategy, we identified the human semaphorin IV gene (H.semaIV, Roche et al., (17)) from within the small cell lung cancer deletion region in proximal 3p21.3. The gene is composed of 18 exons and encodes 785 amino acids having a high degree of similarity with members of the Semaphorin gene family (17). By Northern blot analysis, H.semaIV is expressed in a wide variety of adult tissue mRNAs (17). As initially characterized, Semaphorins (from the Greek sema = sign or signal) are either secreted or transmembrane molecules which regulate the process of nerve growth cone migration. One semaphorin (Collapsin) has been shown to signal through a G-alpha$_i$ protein and is believed to bind to one or more cell surface receptors, although none have yet been identified. Very interestingly, a receptor for netrin, a signaling molecule similar to the semaphorins, has recently been identified and shown to be encoded by the tumor suppressor gene DCC (meeting report). This provides exciting support to the hypothesis that these molecules can have important roles in tumorigenesis. Furthermore, the observation of a wide tissue distribution of H.semaIV transcripts suggests a broader involvement of this gene in the processes it must regulate. We have very preliminary data on the function of H.SemaIV that suggests it may be involved in metastasis. These observations provide a compelling rationale for pursuing this gene and its protein as a candidate tumor suppressor gene in breast cancer.

6.C.4. Immunofluorescence Analysis for H.SemaIV Protein.

We have developed polyclonal antibodies to H.semaIV to permit detection of the protein in cells and tissues. A peptide containing 16 amino acids near the C-terminus of the predicted H.SemaIV protein was synthesized and linked to the carrier protein keyhole limpet hemocyanin (KLH). This material was used to immunize rabbits for production of antibodies specific for H.SemaIV. Blood samples were taken 7 to 10 days after the last injection. Preliminary assessment of antisera titers were determined by immunoblot analyses of 20 to 2000 ng of peptide using

enhanced chemiluminescence. Sera with titers of at least 1:1000 were affinity purified and used as anti- H.SemaIV antibody for immunoprecipitation and immunofluorescence .

Nine breast cancer cell lines have been examined to date. These were fixed using 4% paraformaldehyde at room temperature for 20 minutes or cold Methanol (4°C) for 5 minutes. The cells were washed with phosphate buffered saline (PBS) 3-5 times, then incubated with diluted anti H.SemaIV in 4% BSA; control cells were incubated with 4% BSA at room temperature for 1 hour. After multiple washings with PBS, the cells were incubated with Donkey anti-rabbit antibody conjugated with Texas Red, diluted 1:100 in 4% BSA. The slides were re-washed with PBS, stained with DAPI and stored at 4°C until observed under fluorescence. An example of the results is shown in Fig. 7. Cells from all 9 breast cancer cell lines stained positive with H.SemaIV antibody with cytoplasmic staining. The punctate distribution of fluorescence was suggestive for a Golgi localization and experiments are underway to confirm this. Even though the protein appears to be made in these cell lines, we do not know if it is functional or whether or not the signaling pathway that it is a component of is functional.

6.C.5 Significance

The analysis of this region is going, both at the mRNA level and at the protein level for candidate genes. Very preliminary data suggest a deletion in one of the breast cancer cell lines and we believe that H.SemaIV may have a role in metastasis although we have not observed any alterations of this gene so far. As discussed in the following section, we have actually identified two homozygous deletion regions in 3p21. Which, if any or both, is involved in breast cancer remains to be determined. Our experiments should clarify this important question. These experiments to date address primarily aims 2 and 3 for this region.

6.D. Involvement of 3p21.33 Region in Breast Carcinoma

6.D.1. Development of STS and Polymorphic Markers in Deletion Region.

Chromosomal band 3p21 harbors a second homozygous deletion region, originally described by Yamakawa et al., (40), who showed that the region was missing in 5/36 SCLC cell lines. They went on to show that the minimal deletion spanned approximately 800 kb and was covered by YAC 936C1. We demonstrated that the deletion target was separated from the H.SemaIV deletion site by 10 to 15 megabases of DNA (17). In the course of mapping this region more thoroughly, we discovered that one of our markers, Mbo16E2, was present in this YAC and that it detected homozygous deletions in two uncultured lung tumors (17). This represented the first documented 3p homozygous deletion occurring in a direct tumor, and given the frequent LOH involving 3p21 should now represent a high priority region.

A lambda phage library was constructed from YAC 936C1 in vector λFIXII and 380 human sub-clones (representing the YAC insert approximately 5 times) were isolated by screening with human repetitive DNA probes. The clones were grouped into four sets based upon hybridization results with *Mlu*I restriction fragments from 936C1. Eighty human clones were grown for DNA and sequence analysis was performed from both ends of the clone inserts. This type of sequence sampling is very useful for the development of new PCR amplifiable markers (sequence tagged sites or STSs) and can detect a surprising number of genes when searched against GenBank. One hundred-sixty sequences (about 80 kb or 8 to 10% of the YAC) have been obtained for use as STSs. Useful primer pairs have been obtained from 51. Testing of these primer pairs is underway to ensure that they correctly amplify chromosome 3 specific DNA; of the first 27 tested, 23 yielded the expected product. We are also developing a number of polymorphic markers from this library. Our next step is application of these markers to breast tumor cell line DNAs and direct tumors. Dr. Siggurdur Ingvarsson (Dept. Pathology and Cell Biology, University of Iceland, Reykjavik) has

agreed to analyse a number of these novel polymorphic markers on his collection of breast tumor samples.

We are also building a set of overlapping phage sub-clones around Mbo16E2 since this marker was found homozygously deleted in two uncultured tumors (17). Mbo16E2 identified a single positive clone from the library suggesting that it represents the very end of the YAC, a result consistent with the map position of the locus (41). In order to identify additional cloned DNA from this region, a screen of the chromosome 3-specific cosmid library was undertaken with Mbo16E2 which detected 8 positive independent clones. We are in the process of screening a human PAC library for additional clones.

6.D.2. Sample Sequencing and Predicted Genes.

All 160 sequences were checked against the GenBank non-redundant (NR) and expressed sequence (dbEST) databases using BLAST. Four sequences showed essentially perfect identity to unannotated cDNAs in dbEST. GeneMark exon prediction showed strong coding potential coincident with these cDNA identities. These cDNAs have been ordered from the Research Genetics repository and primer pairs have been made to amplify them. Most interestingly, phage 4G3, which was positive with Mbo16E2, also shows strong sequence homology with one cDNA clone. Thus several genes may have already been identified from within the minimal 3p21.33 deletion region.

6.D.3. Significance

Based on our observed homozygous deletions in direct tumors, as opposed to cell lines, this region seems very likely to encode a tumor suppressor gene. The sequence sampling done to date indicates this area may be gene rich. To our knowledge, this region has not been explored in breast cancer. The reagents under development are designed to rapidly analyze breast carcinoma cell lines for deletions and uncultured tumors for deletions and LOH. This latter group of samples will be more problematic and may require microdissection of tumor cells to eliminate as much of the contaminating stromal material (with normal genomes) as possible. We are excited by the number of potential genes discovered by our sequence sampling efforts and intend to pursue these as candidates as rapidly as possible. We have developed primer pairs for each of these cDNAs that will discriminate between genomic DNA and messenger RNA. Any of these putative genes which appears to be expressed in control tissues and is either not expressed or shows an altered product will be investigated further for mutations. These studies address aims 1 through 3.

6.E. Other potential target genes.

Other target genes will be examined in the upcoming period. These include the type II receptor for tumor growth factor beta (TGFRβII) and the transcription elongation factor A (SII). TGFRβII is particularly interesting since it is a mutation target in GI carcinomas and TGFβ appears to be involved in cell cycle control. The transcription elongation factor A (SII) could be involved by analogy to the known function for the VHL gene product. We will continue to consider new candidate genes as they become available.

7. Conclusions

In each of the previous sections comprising the Body of this report, we have provided a summary of the important results and their significance. Our studies have now explored 4 distinct homozygous deletion regions on 3p. For 3p14, we have identified deletions and rearrangements in breast cancers and have provided strong evidence that a postulated tumor suppressor gene, FHIT, is not the target of these alterations. Additional genes in this region are under investigation. We are also exploring the alternative hypothesis that genomic instability, perhaps mediated by p53

mutation, is responsible for the observed deletions. The reintroduction of YACs is ongoing to provide a functional assessment for tumor suppressor activity.

Similar studies are ongoing for 3 additional homozygous deletion regions (3p12-13, prox. 3p21.3 and distal 3p21.3). For one candidate gene in the proximal 3p21.3 deletion, we have prepared antibodies and are using this reagent for mutation screening. Preliminary evidence indicates that the role of H.SemaIV may be related to metastasis. Several addition genes in this region are undergoing analysis by RT-PCR. For the distal 3p21.3 deletion region, the high frequency of 3p21 LOH in breast cancers combined with our detection of homozygous deletion in direct tumors, indicates that this region must be thoroughly investigated. With our development of a high density of STSs along with the discovery of multiple cDNAs, this objective is being achieved. For the 3p12-13 region, we have tentatively excluded much of the initial U2020 deletion region and have now identified the most likely subsegment for more intense analysis.

8. REFERENCES.

1. Cohen, A. J., Li, F. P., Berg, S., Marchetto, D. J., Tsai, S., Jacobs, S. C. and Brown, R. S. (1979) Hereditary renal-cell carcinoma associated with a chromosomal translocation. *N. Engl. J. Med.*, **301**, 592-595.

2. Boldog, F., Gemmill, R. M., West, J., Robinson, M., Robinson, L., Li, E., Roche, J., Todd, S., Lundstrom, R., Jacobson, J., Klinger, H. and Drabkin, H. A. (1996) 3p14 homozygous deletions and sequence analysis of FRA3B. *Hum. Mol. Genet.*, **Submitted**,

3. Yin, Y., Tainsky, M. A., Bischoff, F. Z., Strong, L. C. and Wahl, G. M. (1992) Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. *Cell*, **70**, 937-948.

4. Livingstone, L. R., White, A., Sprouse, J., Livanos, E., Jacks, T. and Tlsty, T. D. (1992) Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. *Cell*, **70**, 923-935.

5. Lowe, S. W., Ruley, H. E., Jacks, T. and Housman, D. E. (1993) p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. *Cell.*, **74**, 957-967.

6. Sato, T., Akiyama, F., Sakamoto, G., Kasumi, F. and Nakamura, Y. (1991) Accumulation of genetic alterations and progression of primary breast cancer.*Cancer. Res.*, **51**, 5794-579.

7. Bergthorsson, J. T., Eiriksdottir, G., Barkardottir, R. B., Egilsson, V., Arason, A. and Ingvarsson, S. (1995) Linkage analysis and allelic imbalance in human breast cancer kindreds using microsatellite markers from the short arm of chromosome 3. *Hum. Genet*, **96**, 437-443.

8. Pandis, N., Jin, Y., Limon, J., Bardi, G., Idvall, I., Mandahl, N., Mitelman, F. and Heim, S. (1993) Interstitial deletion of the short arm of chromosome 3 as a primary chromosome abnormality in carcinomas of the breast. *Genes. Chromosom. Cancer.*, **6**, 151-155.

9. Chen, L. C., Matsumura, K., Deng, G., Kurisu, W., Ljung, B. M., Lerman, M. I., Waldman, F. M. and Smith, H. S. (1994) Deletion of two separate regions on chromosome 3p in breast cancers. *Cancer Res.*, **54**, 3021-3024.

10. Hainsworth, P. J., Raphael, K. L., Stillwell, R. G., Bennett, R. C. and Garson, O. M. (1991) Cytogenetic features of twenty-six primary breast cancers. *Cancer. Genet. Cytogenet.*, **53**, 205-218.

11. Ali, I. U., Lidereau, R. and Callahan, R. (1989) Presence of two members of c-erbA receptor gene family (c-erbA beta and c-erbA2) in smallest region of somatic homozygosity on chromosome 3p21-p25 in human breast carcinoma. *J. Natl. Cancer. Inst.*, **81**, 1815-1820.

12. Devilee, P., van den Broek, M., Kuipers-Dijkshoorn, N., Kolluri, R., Khan, P. M., Pearson, P. L. and Cornelisse, C. J. (1989) At least four different chromosomal regions are involved in loss of heterozygosity in human breast carcinoma. *Genomics.*, **5**, 554-560.

13. Teixeira, M. R., Pandis, N., Bardi, G., Andersen, J. A. and Heim, S. (1996) Karyotypic comparisons of multiple tumorous and macroscopically normal surrounding tissue samples from patients with breast cancer. *Cancer Res.*, **56**, 855-859.

14. Lu, Y. J., Xiao, S., Yan, Y. S., Fu, S. B., Liu, Q. Z. and Li, P. (1993) Direct chromosome analysis of 50 primary breast carcinomas. *Cancer Genet Cytogenet.*, **69**, 91-99.

15. Buchhagen, D. L., Qiu, L. and Etkind, P. (1994) Homozygous deletion, rearrangement and hypermethylation implicate chromosome region 3p14.3-3p21.3 in sporadic breast-cancer development. *Int. J Cancer*, **57**, 473-479.

16. Gemmill, R. M., Varella-Garcia, M., Smith, D. I., Erickson, P., Golembieski, W., Miller, Y., Coyle-Morris, J., Tommerup, N. and Drabkin, H. A. (1991) A 2.5-Mb physical map within 3p21.1 spans the breakpoint associated with Greig cephalopolysyndactyly syndrome. *Genomics.*, **11**, 93-102.

17. Roche, J., Boldog, F., Robinson, M., Robinson, L., Varella-Garcia, M., Swanton, M., Waggoner, B., Fishel, R., Franklin, W., Gemmill, R. and Drabkin, H. (1996) Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. *Oncogene.*, **12**, 1289-1297.

18. Yokota, J., Tsukada, Y., Nakajima, T., Gotoh, M., Shimosato, Y., Mori, N., Tsunokawa, Y., Sugimura, T. and Terada, M. (1989) Loss of heterozygosity on the short arm of chromosome 3 in carcinoma of the uterine cervix. *Cancer. Res.*, **49**, 3598-3601.

19. Glover, T. W., Berger, C., Coyle, J. and Echo, B. (1984) DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum. Genet.*, **67**, 136-142.

20. Ohta, M., Inoue, H., Cotticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C. M. and Huebner, K. (1996) The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell*, **84**, 587-597.

21. Sozzi, G., Veronese, M. L., Negrini, M., Baffa, R., Cotticelli, M. G., Inoue, H., Pilotti, S., De Gregorio, L., Pastorino, U., Pierotti, M. A., Ohta, M., Huebner, K. and Croce, C. M. (1996) The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell*, **85**, 17-26.

22. Virgilio, L., Shuster, M., Gollin, S. M., Veronese, M. L., Ohta, M., Huebner, K. and Croce, C. M. (1996) FHIT gene alterations in head and neck squamous cell carcinomas. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 9770-9775.

23. Thiagalingam, S., Lisitsyn, N. A., Hamaguchi, M., Wigler, M. H., Willson, J. K. V., Markowitz, S. D., Leach, F. S., Kinzler, K. W. and Vogelstein, B. (1996) Evaluation of the FHIT gene in colorectal cancers. *Cancer Res.*, **56**, 2936-2939.

24. Milde-Langosch, K., Albrecht, K., Joram, S., Schlechte, H., Giessing, M. and Loning, T. (1995) Presence and persistence of HPV infection and p53 mutation in cancer of the cervix uteri and the vulva. *Int. J. Cancer.*, **63**, 639-645.

25. Renault, B., Calistri, D., Buonsanti, G., Nanni, O., Amadori, D. and Ranzani, G. N. (1996) Microsatellite instability and mutations of p53 and TGF-beta RII genes in gastric cancer. *Hum. Genet.*, **In Press**,

26. Wu, C., Akiyama, Y., Imai, K., Miyake, S., Nagasaki, H., Oto, M., Okabe, S., Iwama, T., Mitamura, K. and Masumitsu, H. (1994) DNA alterations in cells from hereditary non-polyposis colorectal cancer patients. *Oncogene.*, **9**, 991-994.

27. Pintado, E., de Diego, Y., Hmadcha, A., Carrasco, M., Sierra, J. and Lucas, M. (1995) Instability of the CGG repeat at the FRAXA locus and variable phenotypic expression in a large fragile X pedigree. *J Med. Genet*, **32**, 907-908.

28. Nancarrow, J. K., Holman, K., Mangelsdorf, M., Hori, T., Denton, M., Sutherland, G. R. and Richards, R. I. (1995) Molecular basis of p(CCG)n repeat instability at the FRA16A fragile site locus. *Hum. Mol. Genet*, **4**, 367-372.

29. Bouffler, S., Silver, A., Papworth, D., Coates, J. and Cox, R. (1993) Murine radiation myeloid leukaemogenesis: relationship between interstitial telomere-like sequences and chromosome 2 fragile sites. *Genes. Chromosomes. Cancer.*, **6**, 98-106.

30. Cohen, S. and Lavi, S. (1996) Induction of circles of heterogeneous sizes in carcinogen-treated cells: two-dimensional gel analysis of circular DNA molecules. *Mol. Cell. Biol.*, **16**, 2002-2014.

31. Misra, R., Shih, A., Rush, M., Wong, E. and Schmid, C. W. (1987) Cloned extrachromosomal circular DNA copies of the human transposable element THE-1 are related predominantly to a single type of family member. *J. Mol. Biol.*, **196**, 233-243.

32. Motejlek, K., Schindler, D., Assum, G. and Krone, W. (1993) Increased amount and contour length distribution of small polydisperse circular DNA (spcDNA) in Fanconi anemia. *Mutat. Res.*, **293**, 205-214.

33. Sunnerhagen, P., Sjoberg, R. M. and Bjursell, G. (1989) Increase of extrachromosomal circular DNA in mouse 3T6 cells on perturbation of DNA synthesis: implications for gene amplification. *Somat. Cell. Mol. Genet.*, **15**, 61-70.

34. Rabbitts, P., Bergh, J., Douglas, J., Collins, F. and Waters, J. (1990) A submicroscopic homozygous deletion at the D3S3 locus in a cell line isolated from a small cell lung carcinoma. *Genes. Chromosom. Cancer.*, **2**, 231-238.

35. Todd, S., Roche, J., Hahner, L., Bolin, R., Drabkin, H. A. and Gemmill, R. M. (1995) YAC contigs covering an 8 megabase region of 3p deleted in the small cell lung cancer cell line U2020. *Genomics.*, **25**, 19-28.

36. Todd, S., Franklin, W. A., Varella-Garcia, M., Kennedy, T., Hilliker, C. E., Hahner, L., Anderson, M., Wiest, J., Drabkin, H. A. and Gemmill, R. M. (1996) Homozygous deletions of human chromosome 3p in lung tumors. *Cancer Res.*, **Submitted**,

37. Sanchez, Y., el-Naggar, A., Pathak, S. and Killary, A. M. (1994) A tumor suppressor locus within 3p14-p12 mediates rapid cell death of renal cell carcinoma in vivo. *Proc. Natl. Acad. Sci. U. S. A.*, **91**, 3383-3387.

38. Killary, A. M., Wolf, M. E., Giambernardi, T. A. and Naylor, S. L. (1992) Definition of a tumor suppressor locus within human chromosome 3p21-p22. *Proc. Natl. Acad. Sci. U. S. A.*, **89**, 10877-10881.

39. Wei, M. H., Latif, F., Bader, S., Kashuba, V., Chen, J. Y., Duh, F. M., Sekido, Y., Lee, C. C., Geil, L., Kuzmin, I., Zabarovsky, E., Klein, G., Zbar, B., Minna, J. D. and Lerman, M. I. (1996) Construction of a 600-kilobase cosmid clone contig and generation of a transcriptional map surrounding the lung cancer tumor suppressor gene (TSG) locus on human chromosome 3p21.3: progress toward the isolation of a lung cancer TSG. *Cancer Res.*, **56**, 1487-1492.

40. Yamakawa, K., Takahashi, T., Horio, Y., Murata, Y., Takahashi, E., Hibi, K., Yokoyama, S., Ueda, R. and Nakamura, Y. (1993) Frequent homozygous deletions in lung cancer cell lines detected by a DNA marker located at 3p21.3-p22. *Oncogene.*, **8**, 327-330.

41. Gemmill, R. M., Chumakov, I., Scott, P., Waggoner, B., Rigault, P., Cypser, J., Chen, Q., Weissenbach, J., Gardiner, K., Wang, H., Pekarski, Y., Le Gall, I., Le Paslier, D., Guillon, S., Li, E., Robinson, L., Hahner, L., Todd, S., Cohen, D. and Drabkin, H. A. (1995) An second-generation YAC-contig map for human chromosome 3. *Nature*, **377S**, 299-319.

9. APPENDIX.

Figure Legends:

Figure 1. Map of clone contig and deletions in 3p14.2

A. Schematic showing position of the homozygous deletion region with respect to YACs 850A6, 65E7 and 74B2. Positions of MluI and selected XhoI restriction sites are indicated. XhoI sites were not determined for YAC 850A6.

B. Cosmid and lambda clone contig spanning the homozygous deletion region within 3p14. Solid horizontal lines represent cosmid (c) and lambda (λ) inserts, as indicated, along with cleavage sites for EcoRI (short vertical bars), XhoI (X) and SalI (S) which were mapped within the central 170 kb. The MluI site corresponding to the site in YAC 74B2 is present in c55D12 and c84F12, although we have not determined its precise location. Positions of selected markers, breakpoints and integration sites for HPV-16 and pSV2neo are indicated across the top. Parts B and C of this figure are drawn to the same scale and positions correspond exactly between the two.

C. Homozygous deletions in tumors and normal genomic DNAs. Horizontal lines denote the extent of deletion in the indicated cell lines. Parallel lines at the ends of each deletion indicate where precise boundaries were not determined. HeLa and MDA231 contain discontinuous deletions indicated by interrupted lines. Exon 5 and the direction of transcription for the FHIT gene are indicated at the bottom.

Figure 2. RT-PCR analysis of FHIT expression in tumor and normal cells.

Panel A. RT-PCR analysis in tumor cell lines. Primers designed to amplify the coding region of FHIT were used to amplify cDNAs prepared from 5 cervical carcinoma lines, one renal carcinoma line (KRC/Y) and an E1A transformed embryonal kidney line (E293), as indicated. Normal products of 638 bp (arrow) were observed in three cell lines, including CC19 with a 40 kb homozygous deletion. A normal sized product was also observed in MS751, although it is amplified weakly.

Panel B. Evidence is presented for alternative splicing of FHIT mRNA in normal tissues. RT-PCR analysis of FHIT expression in fetal brain, adult kidney, E293 cells (whole cells and cytoplasmic fraction), lung tumor A549 and kidney carcinoma cell line KRC/Y were performed as in panel A. Both normal sized products (arrow marked N) and smaller products (left arrow) were observed.

Figure 3. Summary of sequence features observed in FRA3B.

Assembled sequence data totaling 110,435 basepairs. GeneMark and GRAIL2 predicted exons are indicated (filled boxes = direct strand; open boxes = complementary strand). Repetitive sequences homologous to L1 and Alu elements are designated by boxes filled with diagonal lines and shading, respectively. Other repetitive elements are labeled. Unique features of the sequence are indicated on the lowest row.

Figure 4. FISH analysis of A9 cells transfected with YAC 74B2.

Sub-clones of A9 cells transfected with the neomycin resistant form of YAC 74B2 were fixed in methanol/acetic acid and prepared for cytogenetic analysis by spotting onto glass slides. Cosmids c31E1 and c55D12, which are located near the ends of the insert in 74B2g, were labelled with biotin and digoxigenin by nick translation. Following hybridization, the bound 31E1 probe

was visualized by FITC-conjugated avidin and biotinylated goat anti-avidin antibodies (green), while the digoxigenin labelled probe c55D12 was detected with sheep anti-digoxigenin antibodies conjugated with rhodamine (red). Fluorescence signals were observed in a Zeiss Axioskop fluorescence microscope and captured using a BDS image analysis system. The nucleus shown represents a population of cells that contained both sequences in close proximity suggestive that the entire YAC had been integrated intact into the host genomic DNA.

Figure 5. Example of multiplex PCR analysis for the presence or absence of 3p markers in breast carcinoma cell lines.

DNA samples from 8 breast carcinoma cell lines (as indicated) and one normal human cell line (FS) were amplified using primers for two 3p markers which map within 3p21.31 and 3p13, respectively. Amplifications utilized primers for D3S2968 alone (lanes 1), D3S1210 alone (lanes 3) or a mixture of both (lanes 2). The multiplexed amplifications (lanes 2) can be seen to be a sum of the products in lanes 1 and 3 for each sample. The products were separated on an agarose gel and visualized by staining with ethidium bromide. Neither marker was deleted.

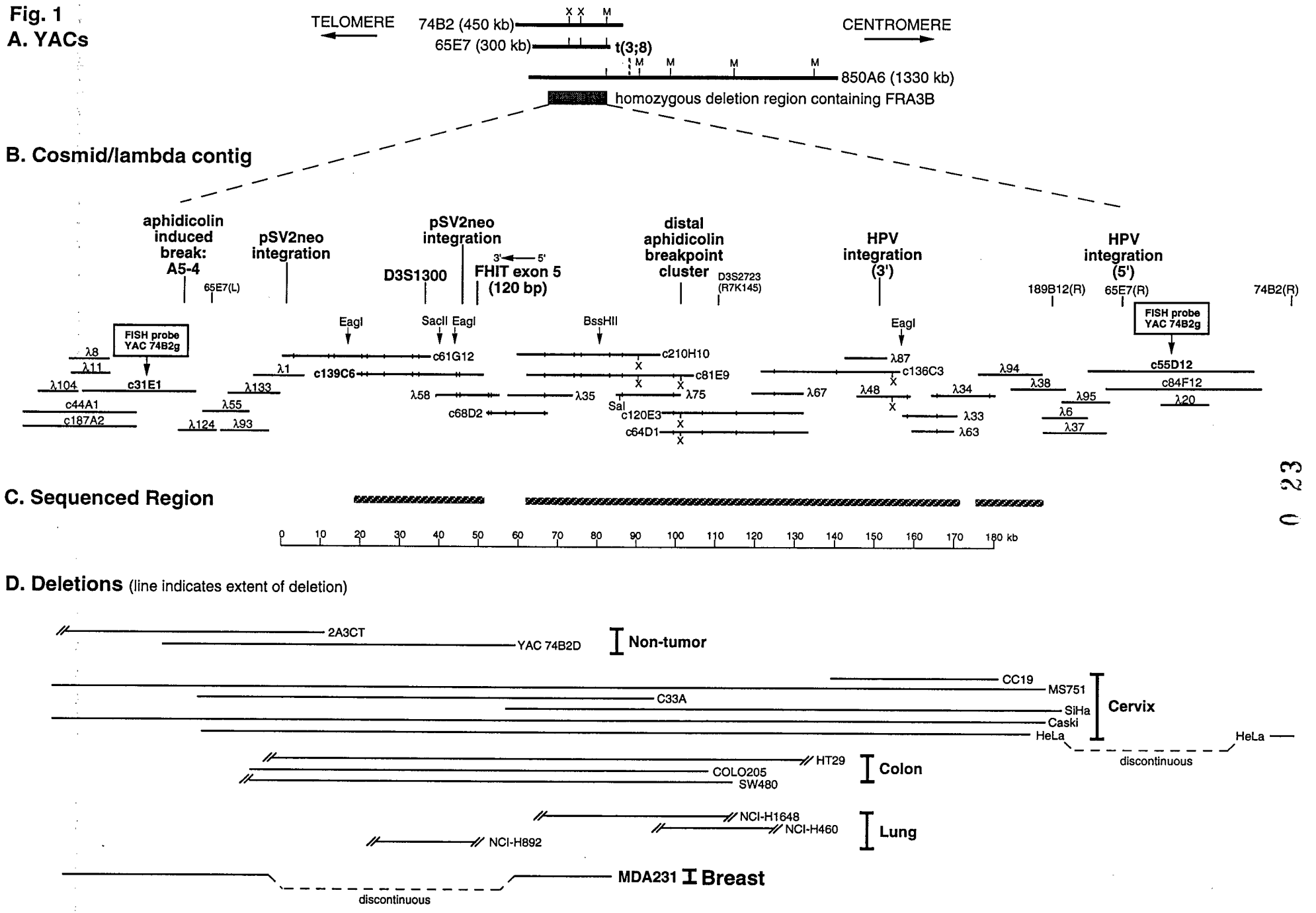
Figure 6. RT-PCR analysis of 3p21.31 genes in breast carcinoma cell lines.

Primers specific for 6 genes (PL6, C02912, LuCa 2, Interferon related, GNATI and LuCa 15) were used to amplify RT-PCR products prepared from 9 breast carcinoma cell lines. Chromosome 3-specific DNA was used as a genomic DNA control. In each panel, the arrow head marks the position of the corresponding genomic DNA product which is larger than the product generated from messenger RNA, as expected. The arrow marks the gene specific product from the cDNA.

Figure 7. Immunohistochemical detection of H.SemaIV in breast cancer cell lines.

Breast cancer cell lines HTB-122 and MDA-231were fixed and incubated with anti H.SemaIV antibody (A and C) and with 4% BSA (B and D) for 1 hour, then the cells were incubated with donkey anti-rabbit IgG conjugated with Texas red. The resulting slides were subjected to immunofluorescence microscopy and captured on a BDS system.

Fig. 1


A. YACs
TELOMERE
74B2 (450 kb)
65E7 (300 kb)
t(3;8)
CENTROMERE
850A6 (1330 kb)
homozygous deletion region containing FRA3B
B. Cosmid/lambda contig
aphidicolin induced break: A5-4
65E7(L)
pSV2neo integration
D3S1300
pSV2neo integration
FHIT exon 5 (120 bp)
distal aphidicolin breakpoint cluster
D3S2723 (R7K145)
HPV integration (3')
HPV integration (5')
189B12(R)
65E7(R)
74B2(R)
FISH probe YAC 74B2g
EagI
SacII
BssHII
λ8
λ11
λ104
c31E1
c44A1
c187A2
λ124
λ93
λ55
λ133
λ1
c139C6
c61G12
λ58
c68D2
λ35
c210H10
c81E9
λ75
Sal
c120E3
c64D1
λ67
λ87
c136C3
λ48
λ34
λ33
λ63
λ94
λ38
λ95
λ6
λ37
c55D12
c84F12
λ20
C. Sequenced Region
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 kb
D. Deletions (line indicates extent of deletion)
2A3CT
YAC 74B2D
Non-tumor
CC19
MS751
C33A
SiHa
Caskl
HeLa
Cervix
HeLa
discontinuous
HT29
COLO205
SW480
Colon
NCI-H1648
NCI-H460
NCI-H892
Lung
MDA231
Breast
discontinuous

C 23

Figure 2. RT-PCR analysis of FHIT

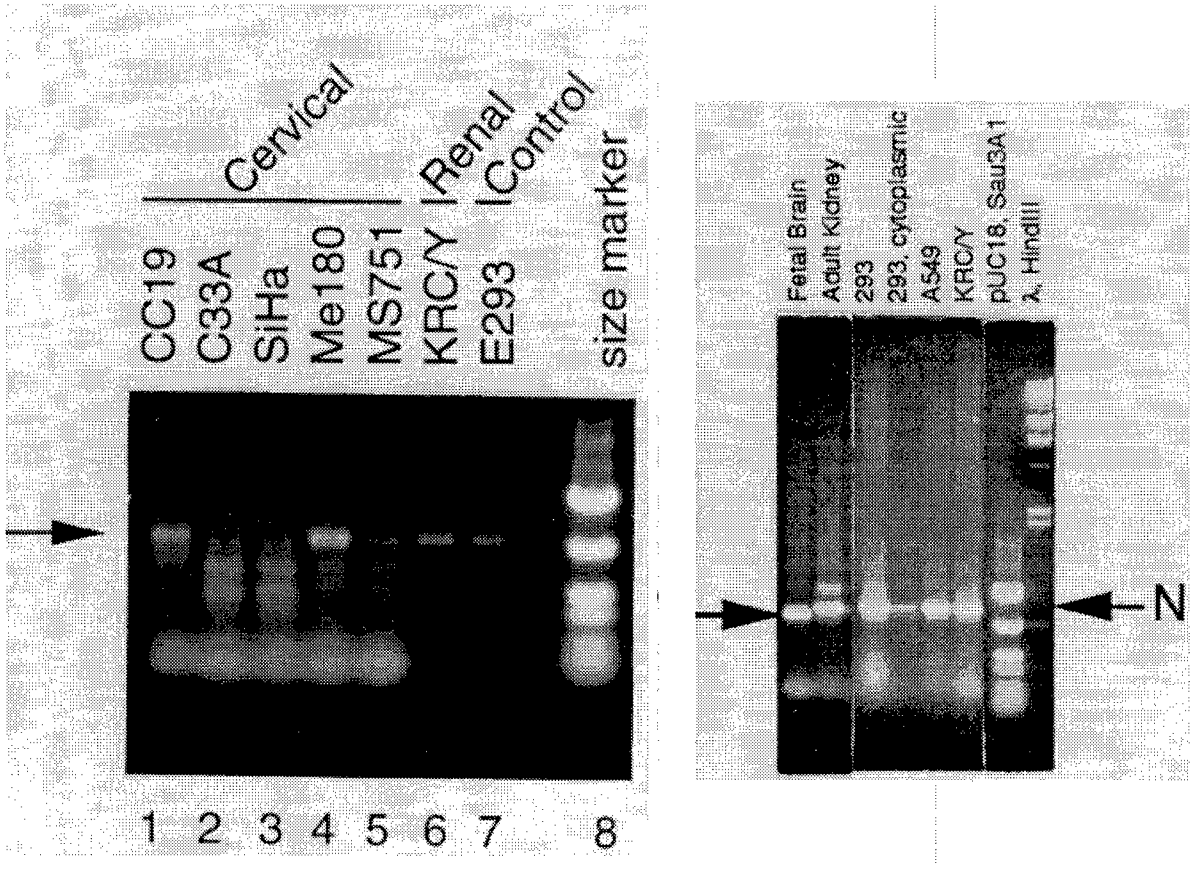


Figure 3

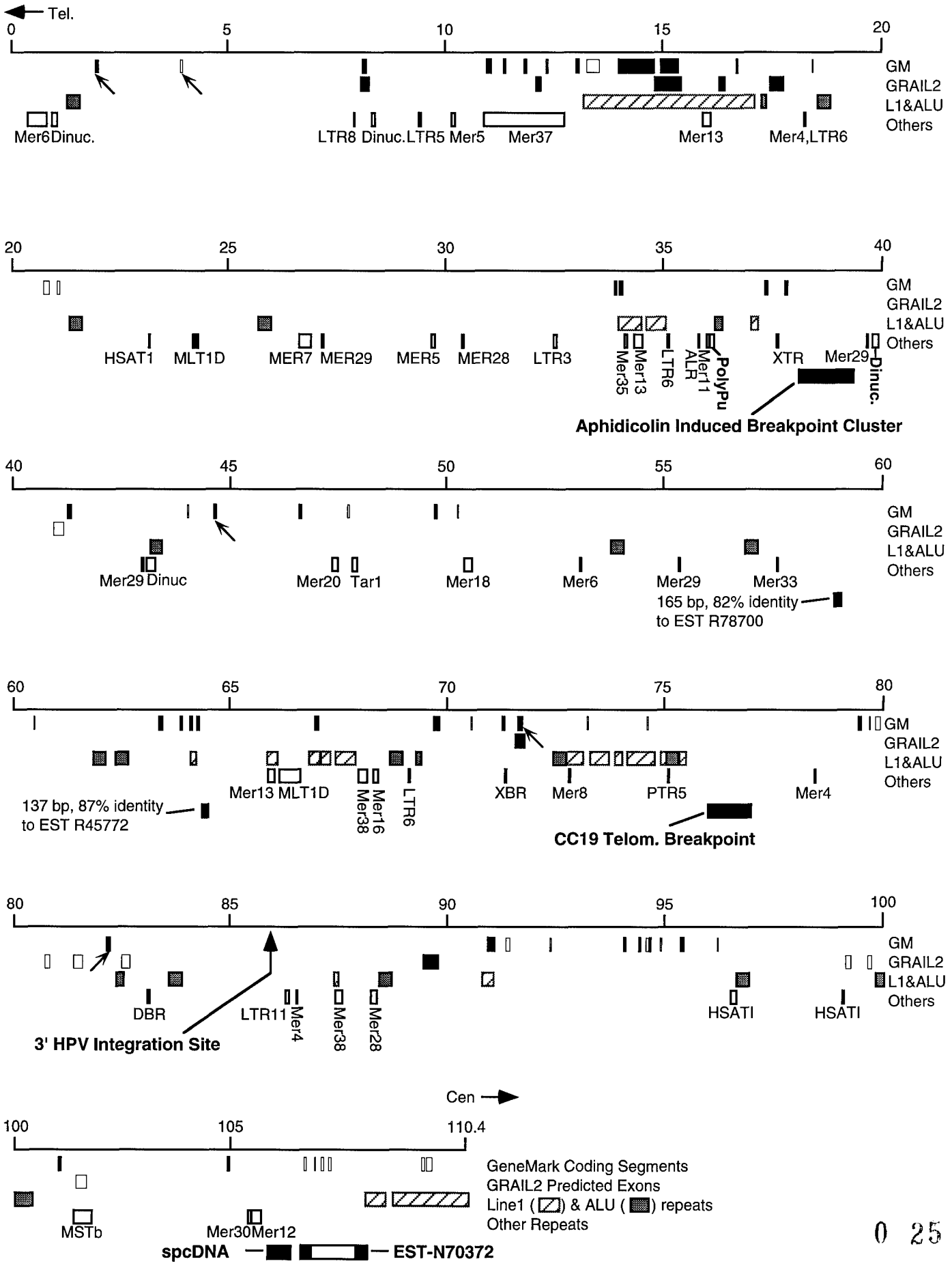

Tel.
0
5
10
15
20
GM
GRAIL2
L1&ALU
Others
Mer6Dinuc.
LTR8 Dinuc.LTR5 Mer5
Mer37
Mer13
Mer4,LTR6
20
25
30
35
40
HSAT1
MLT1D
MER7 MER29
MER5
MER28
LTR3
Mer35
Mer13
LTR6
ALR
Mer11
PolyPu
XTR
Mer29
Dinuc.
Aphidicolin Induced Breakpoint Cluster
40
45
50
55
60
Mer29 Dinuc
Mer20 Tar1
Mer18
Mer6
Mer29
Mer33
165 bp, 82% identity
to EST R78700
60
65
70
75
80
137 bp, 87% identity
to EST R45772
Mer13 MLT1D
Mer38
Mer16
LTR6
XBR
Mer8
PTR5
Mer4
CC19 Telom. Breakpoint
80
85
90
95
100
DBR
LTR11
Mer4
Mer38
Mer28
3' HPV Integration Site
HSATI
HSATI
Cen
100
105
110.4
MSTb
Mer30Mer12
spcDNA
EST-N70372
GeneMark Coding Segments
GRAIL2 Predicted Exons
Line1 ( ) & ALU ( ) repeats
Other Repeats

Figure 4

Figure 5.

PCR analysis for 3p markers in breast carcinoma cell lines.

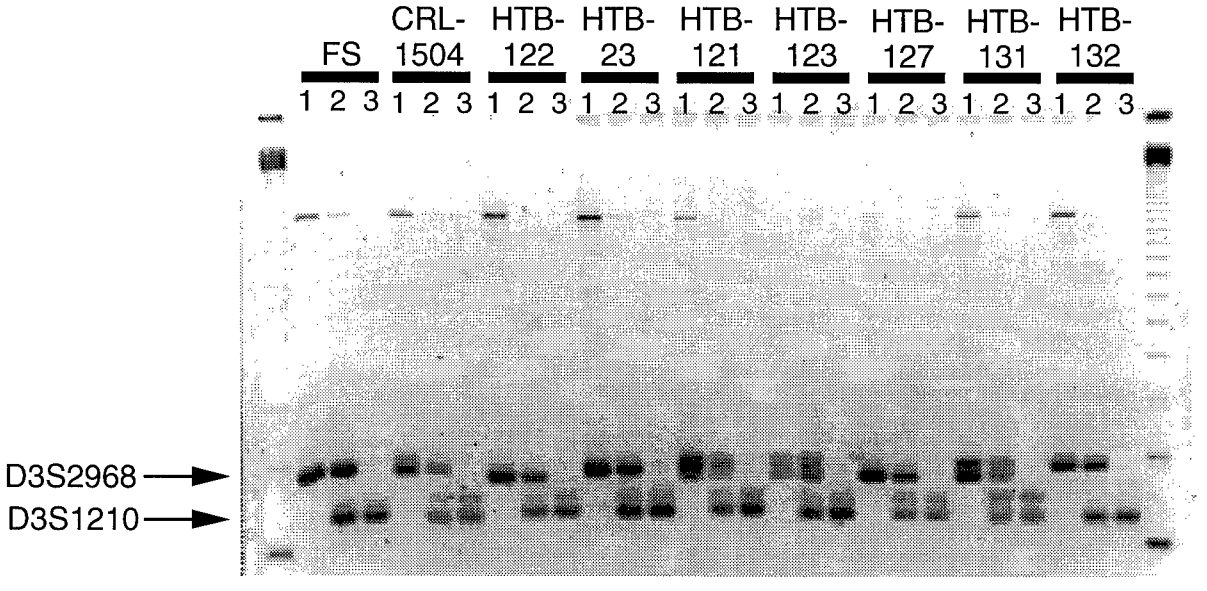
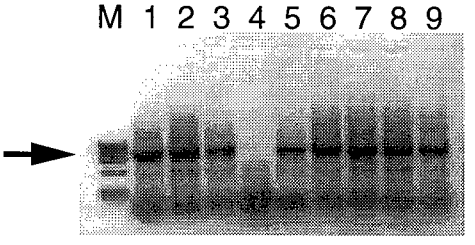
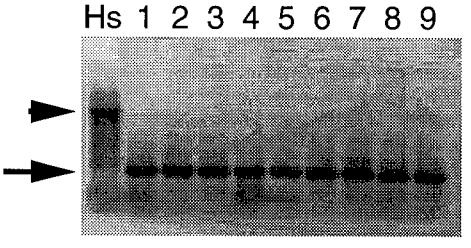


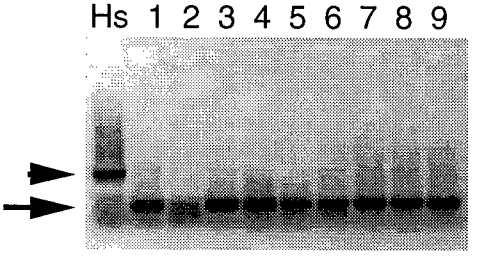
Figure 6: RT-PCR Analysis of 3p21.31 Genes in Breast Carcinoma Cell Lines



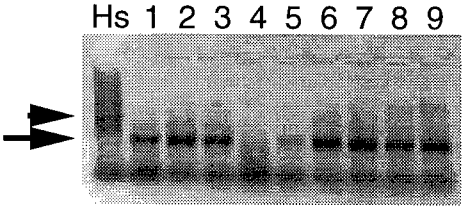
GAPDH Control



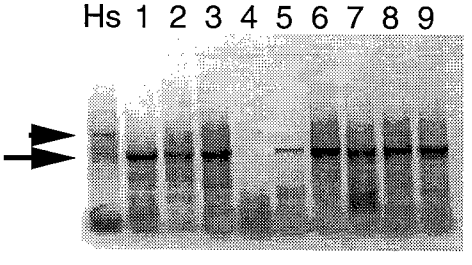
PL6



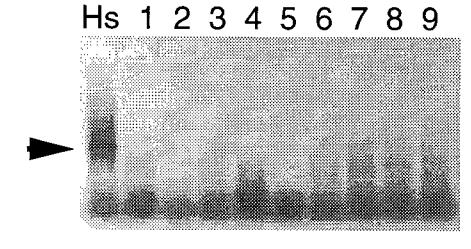
C02912



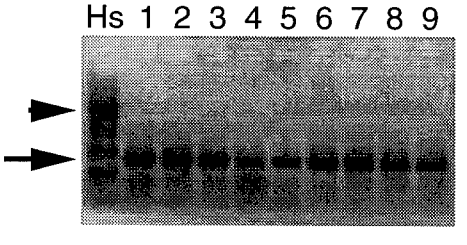
LuCa 2



Interferon related



Rod Transducin (GNATI)



LuCa 15

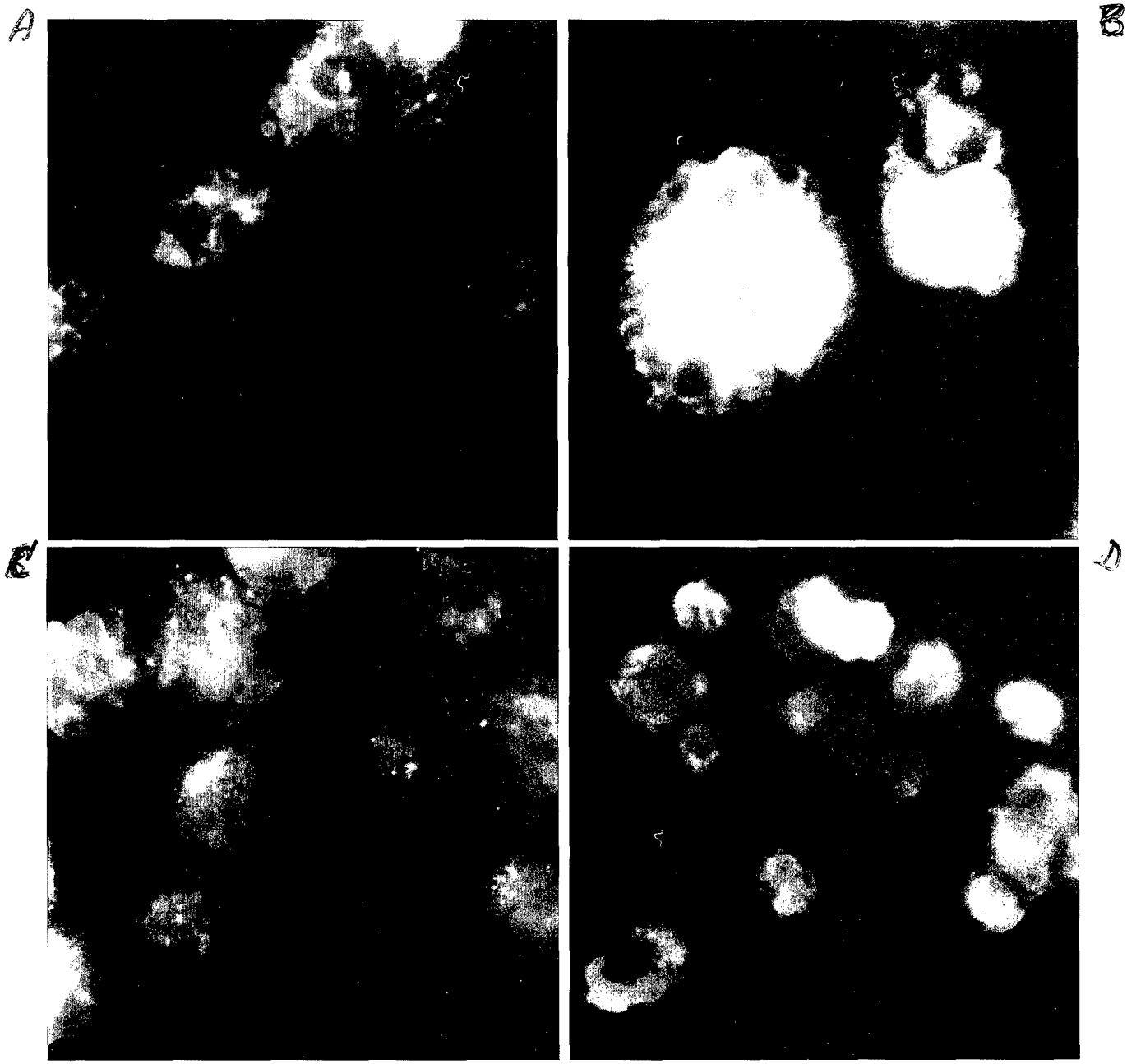


Figure 7

UNPUBLISHED DATA

3p14 HOMOZYGOUS DELETIONS AND SEQUENCE ANALYSIS OF FRA3B

Ferenc Boldog[1*], Robert M. Gemmill[1*], James West[1], Misi Robinson[1], Linda Robinson[1], Efang Li[1], Joelle Roche[1,¶], Sean Todd[1], Ron Lundstrom[2], Jan Jacobson[1], Harold Klinger[3] and Harry A. Drabkin[1†]

[1]Division of Medical Oncology, University of Colorado Health Sciences Center, 4200 E. 9th Ave., Denver, CO 80262; [2]Eleanor Roosevelt Institute, 1899 Gaylord, Denver, Colorado 80206; [3]Dept of Human Genetics, Albert Einstein College of Medicine, 1300 Morris Park Ave., NY, NY 10461; *represents equal contributions; †to whom reprints should be addressed

Address correspondence to: Dr. Harry Drabkin, Division of Medical Oncology, Box B-171, University of Colorado Health Sciences Center, 4200 E. 9th Ave., Denver, CO 80262. Fax 303-270-8825; phone 303-270-4759; e-mail drabkin_h@defiance.uchsc.edu

¶present address: CNRS URA1172, IBMIG, Univ. de Poitiers, 40 Av. du Recteur Pineau, 86022 Poitiers Cédex, France

UNPUBLISHED DATA

Abstract

Loss of heterozygosity (LOH) involving 3p occurs in many carcinomas but is complicated by the identification of 4 distinct homozygous deletion regions. One putative target, 3p14.2, contains the common fragile site, FRA3B, a hereditary renal carcinoma associated 3;8 translocation and the candidate tumor suppressor gene, FHIT. Using a ~300 kb cosmid/lambda contig, we identified homozygous deletions in cervix, breast, lung and colorectal carcinoma cell lines. The smallest deletion (CC19) was shown not to involve FHIT coding exons and no DNA sequence alterations were present in the transcript. We also detected discontinuous deletions as well as deletions in non-tumor DNAs suggesting that FHIT is not a selective target. Further, we demonstrate that some reported FHIT aberrations represent normal splicing variation. DNA sequence analysis of 110 kb demonstrated that the region is high in A-T content, LINEs and MER repeats, whereas Alu elements are reduced. We note an intriguing similarity in repeat sequence composition between FRA3B and a 152 kb segment from the Fragile-X region. We also identified similarity between a FRA3B segment and an spcDNA. In contrast to the selective loss of a tumor suppressor gene, we propose an alternative hypothesis, that some putative targets including FRA3B may undergo loss as a consequence of genomic instability. This instability is not due to DNA mismatch repair deficiency, but may correlate in part with p53 inactivation.

INTRODUCTION

LOH involving 3p occurs frequently in carcinomas of the lung, kidney, cervix, breast and other epithelial neoplasms (1-7). However, 3p loss is complex, involving at least four distinct homozygously deleted regions (8-12). One of the most frequently lost regions is 3p14, especially in cervical carcinomas (13). This region is also of interest since it contains the site of a hereditary renal carcinoma associated translocation, t(3;8)(p14.2;q24.1) (14), and is the location of the most inducible common fragile site in the genome, FRA3B (15). We previously reported cloning of the 3;8 translocation breakpoint (16) and demonstrated by FISH that FRA3B was further telomeric (17). These studies also suggested that breaks in FRA3B occur over a region instead of a single site. While searching for genes in this region, we identified a homozygous deletion in the cervical carcinoma cell line, HeLa, involving marker D3S1300.

We developed a ~300 kb cosmid/lambda contig within FRA3B containing D3S1300. Probes from the region detected frequent homozygous deletions in cervical, lung, colorectal and breast carcinoma cell lines. Cervical carcinomas, which are associated with papilloma virus infection and p53 inactivation (18), were most frequently deleted and the smallest deletion occurred in cell line CC19. During our investigations, Ohta et al. reported identification of a candidate tumor suppressor gene, FHIT, which spanned the t(3;8) breakpoint and was deleted in various carcinoma cell lines (19). However, our analysis indicates that FHIT is not the target of these deletions. We also observed that 3p14 deletions tend not to occur in tumors with deficiencies in DNA mismatch repair.

DNA sequencing studies were performed to identify features that might provide insight into these breaks. Our results demonstrate that FRA3B is high in A-T content, LINES and MER repeats. In contrast, Alu elements and confirmed genes are reduced. We identified a FRA3B segment highly similar to a reported small polydispersed circular DNA, sequences which are

markedly elevated in damaged or unstable genomes. FRA3B also bears an intriguing overall sequence similarity to the Fragile X region. However, unlike the rare folate-sensitive fragile sites, no triplet repeats nor methylated CpG-island was identified. These overall features may be responsible for, or contribute to, the observed instability of this region.

RESULTS

Identification of Homozygous Deletion Following identification of the HeLa deletion, additional cervical carcinoma lines were examined using probes spanning a ~300 kb cosmid/phage contig from the FRA3B region (Fig. 1A and B). Fig. 2A shows an example using cosmid c136C3. Deletions were evident in DNAs from MS751, SIHA and Caski (lanes 2, 5 and 7) whereas CC19 (lane 6) shows missing and altered bands (arrow). Altogether, homozygous deletions were detected in 7 of 8 lines (87.5%). Similar hybridizations were performed using DNAs from 12 colon tumor lines and deletions or rearrangements were seen in 50% (Fig. 2B). Other homozygous deletions (Fig. 1C) were detected in 3 lung (NCI-H1648, NCI-H460 and NCI-H892) and 1 breast carcinoma (MDA-231) cell lines. The smallest (~40 kb) was in the cervical carcinoma line CC19. This line was further characterized by DNA sequence analysis and FHIT gene expression (see below). Two (HeLa and MDA-231) contained discontinuous deletions which was surprising since a single deletion should have been sufficient to inactivate a tumor suppressor gene. No deletions were detected in 5 renal carcinoma lines (KRC/Y, CAKI-1, CAKI-2, ACHN and KV6).

We also observed deletions in non-tumor derived DNAs. Hybridization with c61G12 to a chromosome 3 hybrid panel revealed a partial deletion in 2A3CT (Fig. 3A). Similarly, a spontaneous 80 kb deletion in YAC 74B2 spanning the region was identified during single clone purification (74B2Δ). Fig. 3B shows missing bands (arrows) in 74B2Δ when hybridized with c31E1. This 80 kb deleted segment encompasses the aphidicolin-induced breakpoint in hybrid A5-

4 (Figs. 3C and 1C), the ends of several tumor deletions (C33A, SIHA, HeLa, COLO-205 and MDA-231) and the more telomeric of two pSV2neo plasmid insertion sites which preferentially integrated into FRA3B after aphidicolin treatment (20,21).

Correlation of Deletions with Microsatellite Instability In cervical and colorectal carcinoma cell lines, we observed a trend of inverse correlation between the presence of a 3p14 deletion and reported microsatellite instability (22-26). However, for some lines we were unable to discern the replication error (RER) status from the literature. Therefore, we subcloned selected lines and tested 10 clones each with up to 6 microsatellite loci (D3S1300, D3S1210, D3S1286, D3S1233 and AFM320yb5). Instability was accepted if new bands appeared from 2 or more loci, although there were no cases of only one alteration. Results for samples where we have information on both DNA mismatch repair and 3p14 deletion are shown in Table I. While the number of RER+ lines is small, both 3p14 deletions and the RER phenotype are usually discordant. One simple conclusion is that defects associated with microsatellite instability alone are not responsible for the observed deletions. We also observed the highest incidence of deletions in cervical carcinoma lines, where p53 alterations appear very common (27,28). This may relate to the reduced frequency of deletions observed in RER+ colorectal tumors (see Discussion).

Relative Position of FHIT To examine the role of the FHIT gene in these deletions, we derived primers (FHIT-5' and FHIT-3') for a one-stage RT-PCR amplification of the entire coding region (exons 5-9). An expected 638 bp product was amplified after 35 cycles from embryonal kidney 293 cells (Fig. 4A, lane 7). This was used as a probe against the contig which showed that only a single FHIT exon, located adjacent to D3S1300, was present (Fig. 1B). Comparison of our map to that described by Ohta et al., (19) allowed us to conclude that this was exon 5 (also confirmed by DNA sequencing, not shown) and demonstrated that several tumor deletions exclusively affected large introns (Fig. 1C). This occurred in CC19 and SIHA (cervix), NCI-H1648 and NCI-H460 (lung) and MDA-231 (breast).

Analysis of FHIT in the tumor lines shown in Fig. 4A demonstrated that 3/6, including CC19, contained normal sized bands (arrow). To further characterize the FHIT product from CC19, 4 independent isolates were sequenced. Each contained only normal sequences comprising coding exons 5-9. Thus, the homozygous deletion had no apparent effect on FHIT mRNA suggesting either a different target gene or unselected genomic instability. Additionally, a faint normal sized product was seen in MS751 (Fig. 4A, lane 5) which by Southern blot contains a homozygous deletion including exon 5. A possible explanation is that the deletion is heterogeneous within the cell population suggesting it occurred during culture.

The cervical carcinoma lines C33A and SIHA contained multiple smaller RT-PCR products with no detectable wild-type product. While C33A had an exon 5 deletion that could explain one smaller band (Fig. 1C), the multiplicity of products suggested alternative splicing. The SIHA deletion does not affect an exon although there may be non-recognized discontinuous or overlapping bi-allelic deletions. Despite these obvious differences, the RT-PCR products appeared identical (Fig. 4A, lanes 2&3). To examine this, we amplified FHIT coding exons from RNA or cDNAs prepared from normal (fetal brain, adult kidney), immortalized (E293) and tumor (A549, KRC/Y) samples (Fig. 4B). While normal FHIT (arrow marked N) was observed in each, a smaller product was also seen which predominated in the normal fetal brain cDNA library. When this was cloned and sequenced, exon 8 containing the conserved histidine triad motif and possible zinc binding site (29) was missing (not shown). This variation had been reported to represent an aberration in squamous cell carcinomas of the head and neck (30). Additional larger products were identified from the adult kidney cDNA library although these have not been characterized. Thus, alternative splicing definitely occurs in normal tissues.

DNA Sequence Analysis Since CC19 contained the smallest deletion, we initiated large scale DNA sequencing studies to identify new genes or structural features that might be responsible for

the genomic instability. Six genome equivalents of sequence were obtained from c81E9, c120E3, c136C3 and λ33 and the data assembled. Gaps were closed using directed primers and a 1.3 kb clone gap between c136C3 and λ33 was closed by PCR amplification using λ48 (Fig. 1B) as template. This resulted in 110.4 kb of contiguous sequence (GenBank U66722) at an accuracy of ~99.6%. Predicted restriction maps from the assembled sequence also matched those determined experimentally. Sequence analysis (shown schematically in Fig. 5) included similarity searches, repetitive sequence identification and exon prediction.

Gene Search Overall, the sequence is AT-rich (61.1%) and very depleted in CG dinucleotides. No identities to known genes were seen. GRAIL2 predicted fifteen exons of which six had moderate (<0.4), five had good (0.4-0.6) and four had excellent (>0.6) scores. Four putative exons were coincident with repeat sequences and two others (positions 8.0 and 71.5 kb) were coincident with Genemark predicted coding segments. Genemark (31) was utilized in order to implement a matrix for higher A-T content regions. While 61 potential coding segments were identified, no significant similarities were observed. Many predicted coding segments clearly occurred within LINE and MER elements and others not directly within repeats nevertheless demonstrated similarity to LINEs. Four putative exons were clustered near position 107 kb, one of which showed perfect identity with a partially sequenced cDNA, EST-N70372. However, Northern analysis failed to identify a transcript from this region (data not shown) and the cDNA sequence included a portion of the LINE element at position 108 kb. Based on end-sequences from the cDNA clone and insert length, the cDNA was co-linear with genomic DNA. Further, the 3' end of the cDNA was coincident with a polyA tract in genomic DNA, all suggesting this represented an unprocessed transcript, or more likely resulted from false priming of contaminating DNA. BLASTN searches determined two additional regions with similarity to non-annotated cDNAs (positions 58.2 kb and 64.3 kb). However, neither showed similarity to known genes nor were directly superimposed on predicted exons. That these sequences were observed adjacent to, rather than superimposed on, predicted exons could be due to conservative prediction algorithms

used which underestimate the extent of many exons. Alternatively, the homologies could be due to infrequent repeats. We note the presence of 5 remaining putative exons, denoted in Figure 5, having high probability scores and which do not overlap repeats.

Repeat Sequence Analysis Analysis with Pythia (32) showed that 20.2% of the sequenced region is comprised of known repeats, a level comparable to other regions chosen for comparison (Table II). However, in this 110.4 kb, LINE and MER elements make up the bulk of repeats. Intriguingly, the repeat composition is very similar to a 152 kb sequence from the Fragile-X region (33) and differs from arbitrarily selected segments in 3p21.3 and 4p16.3 (Table II). In particular, λ33 (from pos. 96 to 110.4 kb) within the CC19 deletion is nearly identical in LINE element composition to the Fragile-X region. Similarly, both Fragile-X and FRA3B have a low level of predicted coding regions. Speculatively, these similarities may influence the observed tendency of both regions to undergo breakage. However, there are obvious differences between the two sites, notably the presence of a triplet repeat and methylated CpG-island in Fragile-X (34) and the common vs. rare nature of FRA3B.

Other features Position 38 kb contains a cluster of terminal deletion breakpoints induced by aphidicolin treatment (GenBank U46001) further linking our sequence to FRA3B. Flanking these aphidicolin breakpoints is a long polypurine tract and two extended variable dinucleotide repeats. The telomeric breakpoint of the CC19 deletion lies within a 1.0 kb region at position 76 kb between L1 and MER/LTR elements. Within the CC19 deletion region, near position 86 kb, is an HPV16 integration site from a cervical carcinoma (35). Of note, this integration was associated with an interstitial deletion (35,36) which we now know would not affect FHIT coding sequences. Interestingly, a very significant similarity ($1e^{-77}$) to a small polydispersed circular (spc) DNA was observed at 106 kb (GenBank X96885). Fig. 6 shows a FASTA alignment over 585 bp. The region is 72.1% identical and contains interspersed blocks having up to 89% identity (e.g., spcDNA bp 121-207). This spcDNA, which appears non-repetitive, was isolated from a tuberous

sclerosis-associated angiofibroma (Ph.D. thesis, I. Hinkel-Schreiner). Characteristics of spcDNAs include derivation from chromosomal sequences (37), association with clustered repeats (such as β-satellites and other clustered elements although a single family member may be predominantly involved (38)) and elevation in conditions associated with genomic instability such as Fanconi's anemia (39). spcDNAs are also increased by DNA damaging agents (37) and inhibitors of DNA and protein synthesis including the fragile site inducer aphidicolin (40). Given the limited number of spcDNAs that have been sequenced, this similarity may be biologically important.

DISCUSSION

Using a ~300 kb cosmid/lambda contig, located approximately 150 kb telomeric to the 3;8 translocation breakpoint, we have identified homozygous deletions in various carcinoma cell lines that overlap the most inducible common fragile site in the genome, FRA3B. From various aphidicolin induced breakpoints and plasmid or viral integration sites (Fig. 1), FRA3B represents a region rather than a single site. Studies by Wilke et al. (36) and Smith et al (41) indicate that some clustering of breakpoints may occur. However, their rearrangements were induced by aphidicolin in a single chromosome 3 containing hybrid and, unlike the interstitial deletions we observed in tumor and non-tumor samples, appear to represent terminal breaks. Where we have accurately defined the boundaries for the carcinoma-associated deletions, one or both are contained within the FRA3B region.

Our initial hypothesis was that the smallest deletion (CC19) would contain elements of a tumor suppressor gene. During our studies, Ohta et al. (19) identified the FHIT gene with reported abnormalities in RT-PCR products. However, FHIT has similarity to a yeast di-adenosine hydrolase which would represent an unexpected function for a tumor suppressor gene. Our results indicate that FHIT is not the target of these deletions. First, the CC19 deletion does not involve

FHIT coding sequences (Fig. 1B&C) and based on RT-PCR and cDNA sequence analysis the coding portion of the FHIT transcript is normal. Second, we have observed deletions in genomic DNAs from non-tumor sources (Fig. 3A). The somatic cell hybrid 2A3CT was derived by spontaneous terminal deletion at 3p21.3 from a normal chromosome 3 hybrid, UCTP2A3 (11). However, 2A3CT also acquired a 3p14 interstitial deletion overlapping an aphidicolin induced breakpoint, a pSV2neo integration into FRA3B (20) and the telomeric borders of several carcinoma associated deletions. We also identified an overlapping deletion in an unselected subclone from YAC 74B2. While neither the hybrid nor YAC 3p14 DNA segments are in a "normal" background, they clearly are unselected from a tumorigenic standpoint. We note that "hotspots" of recombination in human DNA can be maintained in a yeast background (42), thus it is not unreasonable that unstable regions may behave similarly or more so. Third, discontinuous deletions appear common in this region (i.e., HeLa, MDA231), both from our analysis and from that reported by Ohta et al (19). Multiple deletions might be expected if there was no common target gene and if the region was unstable. Fourth, we have observed that FHIT undergoes alternative splicing in normal tissues (Fig. 4B) which explains some previously reported abnormal PCR products (19,30,43). Thiagalingam et al (44) recently reported lack of FHIT involvement in colorectal carcinomas and suggested that PCR artifacts might be responsible for some previously observed alterations (19,43). It also seems likely that many deletions would have been missed in their study since only few markers were tested. One of the possible features that suggested FHIT could be a tumor suppressor gene was that it crossed the hereditary renal carcinoma associated 3;8 breakpoint (19). However, we found no alterations in RT-PCR products from 5 renal carcinoma cell lines. Moreover, Bugert et al. (manuscript submitted) have observed normal FHIT transcripts and no point mutations in a large series of renal cancers. Thus, FHIT does not appear to be involved in renal carcinoma. With respect to other possible target genes, from our sequencing studies we identified a 100% identity to two ESTs from a liver/spleen library. However, we determined that both clones were identical, were not expressed using a commercial Northern blot

(Clontech), were co-linear with genomic DNA containing a polyA tract corresponding to their 3' end, encoded no significant open reading frame and overlapped a partial LINE element.

While other tumor suppressor genes may exist within FRA3B, an alternative possibility is that the deletions are due to primary genomic instability affecting a particularly susceptible region. This hypothesis is consistent with several of the observations reported here including their discontinuity and occurrence in non-tumor cell lines. By using numerous probes, we were able to identify a high frequency of homozygous deletions, especially in cervical carcinomas where p53 inactivation is very common (18). In this regard, it is interesting that we observed an inverse correlation between 3p14 deletions and microsatellite instability (RER+). Importantly, p53 mutations, which have been shown to destabilize the genome (45,46), appear infrequent in RER+ colorectal carcinomas and gastric tumors ((47,48) and P. Cottu, presented at Cancer and the Cell Cycle, Lausanne, Switzerland 1996). Thus, these findings would be consistent with 3p14 deletions resulting from the genomic instability which accompanies p53 inactivation.

What have we learned from the DNA sequence analysis to date ? First, is that the region is high in A-T content with frequent LINEs and MER repeats, and is conversely low in Alu sequences and confirmed genes. In contrast to the reported rare folate-sensitive sites which are associated with expanded CGG repeats (49,50), FRA3B does not contain an expanded triplet repeat or methylated CpG-island. Nor did we identify any telomeric repeats which have been suggested as a possible cause of breaks (51). However, we do note that over the 110.4 kb region there is an overall repeat sequence similarity to 152 kb from Fragile-X. In vitro, expanded CGG repeats have been shown to inhibit DNA replication (52). Whether or not specific sequences within FRA3B may have a similar effect on replication awaits experimental testing. Our discovery of a strong spcDNA homology in FRA3B may therefore not be coincidental. While we do not know if the spcDNA site is a primary cause of FRA3B instability or simply a marker for this property, the DNA sequences reported should provide the means to test this.

ACKNOWLEDGEMENTS This work was supported by DAMD17-94-J-4391 (HAD, RMG) and the DNA sequencing studies were supported by HG000358 (HAD, RMG). We gratefully acknowledge support from the Colorado Cancer League (2531718) to FB. Additional Core support was provided through the University of Colorado Cancer Center (CA 46934)

REFERENCES

1. Ogawa, O., Kakehi, Y., Ogawa, K., Koshiba, M., Sugiyama, T. and Yoshida, O. (1991) Allelic loss at chromosome 3p characterizes clear cell phenotype of renal cell carcinoma. *Cancer. Res.*, **51**, 949-953.

2. Devilee, P., van den Broek, M., Kuipers-Dijkshoorn, N., Kolluri, R., Khan, P. M., Pearson, P. L. and Cornelisse, C. J. (1989) At least four different chromosomal regions are involved in loss of heterozygosity in human breast carcinoma. *Genomics.*, **5**, 554-560.

3. Hibi, K., Takahashi, T., Yamakawa, K., Ueda, R., Sekido, Y., Ariyoshi, Y., Suyama, M., Takagi, H. and Nakamura, Y. (1992) Three distinct regions involved in 3p deletion in human lung cancer. *Oncogene.*, **7**, 445-449.

4. Lothe, R. A., Fossa, S. D., Stenwig, A. E., Nakamura, Y., White, R., Borresen, A. L., Brogger, A., Brresen, A. L. and Brgger, A. (1989) Loss of 3p or 11p alleles is associated with testicular cancer tumors. *Genomics.*, **5**, 134-138.

5. Yang-Feng, T. L., Han, H., Chen, K. C., Li, S. B., Claus, E. B., Carcangiu, M. L., Chambers, S. K., Chambers, J. T. and Schwartz, P. E. (1993) Allelic loss in ovarian cancer.. *Int. J. Cancer.*, **54**, 546-551.

6. Kohno, T., Takayama, H., Hamaguchi, M., Takano, H., Yamaguchi, N., Tsuda, H., Hirohashi, S., Vissing, H., Shimizu, M., Oshimura, M. and et alX, (1993) Deletion mapping of chromosome 3p in human uterine cervical cancer. . *Oncogene.*, **8**, 1825-1832.

7. Maestro, R., Gasparotto, D., Vukosavljevic, T., Barzan, L., Sulfaro, S. and Boiocchi, M. (1993) Three discrete regions of deletion at 3p in head and neck cancers. *Cancer Res.*, **53**, 5775-5779.

8. Yamakawa, K., Takahashi, T., Horio, Y., Murata, Y., Takahashi, E., Hibi, K., Yokoyama, S., Ueda, R. and Nakamura, Y. (1993) Frequent homozygous deletions in lung cancer cell lines detected by a DNA marker located at 3p21.3-p22. *Oncogene.*, **8**, 327-330.

9. Drabkin, H. A., Mendez, M. J., Rabbitts, P. H., Varkony, T., Bergh, J., Schlessinger, J., Erickson, P. and Gemmill, R. M. (1992) Characterization of the submicroscopic deletion in the small- cell lung carcinoma (SCLC) cell line U2020. *Genes. Chromosom. Cancer.*, **5**, 67-74.

10. Rabbitts, P., Bergh, J., Douglas, J., Collins, F. and Waters, J. (1990) A submicroscopic homozygous deletion at the D3S3 locus in a cell line isolated from a small cell lung carcinoma.. *Genes. Chromosom. Cancer.*, **2**, 231-238.

11. Roche, J., Boldog, F., Robinson, M., Robinson, L., Varella-Garcia, M., Swanton, M., Waggoner, B., Fishel, R., Franklin, W., Gemmill, R. and Drabkin, H. (1996) Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. *Oncogene.*, **12**, 1289-1297.

12. Lisitsyn, N. A., Lisitsina, N. M., Dalbagni, G., Barker, P., Sanchez, C. A., Gnarra, J., Linehan, W. M., Reid, B. J. and Wigler, M. H. (1995) Comparative genomic analysis of tumors: Detection of DNA losses and amplification. *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 151-155.

13. Yokota, J., Tsukada, Y., Nakajima, T., Gotoh, M., Shimosato, Y., Mori, N., Tsunokawa, Y., Sugimura, T. and Terada, M. (1989) Loss of heterozygosity on the short arm of chromosome 3 in carcinoma of the uterine cervix. *Cancer Res.*, **49**, 3598-3601.

14. Cohen, A. J., Li, F. P., Berg, S., Marchetto, D. J., Tsai, S., Jacobs, S. C. and Brown, R. S. (1979) Hereditary renal-cell carcinoma associated with a chromosomal translocation. . *N. Engl. J. Med.*, **301**, 592-595.

15. Glover, T. W., Berger, C., Coyle, J. and Echo, B. (1984) DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes.. *Hum. Genet.*, **67**, 136-142.

16. Boldog, F. L., Gemmill, R. M., Wilke, C. M., Glover, T. W., Nilsson, A. S., Chandrasekharappa, S. C., Brown, R. S., Li, F. P. and Drabkin, H. A. (1993) Positional cloning of the hereditary renal carcinoma 3;8 chromosome translocation breakpoint. . *Proc. Natl. Acad. Sci. U. S. A.*, **90**, 8509-8513.

17. Wilke, C. M., Guo, S. W., Hall, B. K., Boldog, F., Gemmill, R. M., Chandrasekharappa, S. C., Barcroft, C. L., Drabkin, H. A. and Glover, T. W. (1994) Multicolor FISH mapping of YAC clones in 3p14 and identification of a YAC spanning both FRA3B and the t(3;8) associated with hereditary renal cell carcinoma. *Genomics.*, **22**, 319-326.

18. Milde-Langosch, K., Albrecht, K., Joram, S., Schlechte, H., Giessing, M. and Loning, T. (1995) Presence and persistence of HPV infection and p53 mutation in cancer of the cervix uteri and the vulva. *Int. J. Cancer.*, **63**, 639-645.

19. Ohta, M., Inoue, H., Cotticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C. M. and Huebner, K. (1996) The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell*, **84**, 587-597.

20. Rassool, R. V., Le Beau, M. M., Shen, M. -L., Neilly, M. E., Espinosa, R. III, Ong, S. T., Boldog, F., Drabkin, H., McCarroll, R. and McKeithan, T. W. (1996) Direct cloning of DNA sequences from the common fragile site region at chromosome band 3p14.2. *Genomics*, **In Press**,

21. Rassool, F. V., McKeithan, T. W., Neilly, M. E., van Melle, E., Espinosa, R. and LeBeau, M. M. (1991) Preferential integration of marker DNA into the chromosomal fragile site at 3p14: an approach to cloning fragile sites. *Proc. Natl. Acad. Sci. U. S. A.*, **88**, 6657-6661.

22. Bhattacharyya, N. P., Ganesh, A., Phear, G., Richards, B., Skandalis, A. and MeuthM., (1996) Molecular analysis of mutations in mutator colorectal carcinoma cell lines. *Hum. Mol. Genet.*, **4**, 2057-2064.

23. Boyer, J. C., Umar, A., Risinger, J. I., Lipford, J. R., Kane, M., Yin, S., BarrettJC., , Kolodner, R. D. and Kunkel, T. A. (1995) Microsatellite instability, mismatch repair deficiency, and genetic defects in human cancer cell lines. *Cancer Res.*, **55**, 6063-6070.

24. da Costa, L. T., Liu, B., el-Deiry, W., Hamilton, S. R., Kinzler, K. W., VogelsteinB., , Markowitz, S., Willson, J. K., de la Chapelle, A., Downey, K. M. and et al., (1995) Polymerase delta variants in RER colorectal tumours [letter]. *Nat. Genet.*, **9**, 10-11.

25. Umar, A., Boyer, J. C., Thomas, D. C., Nguyen, D. C., Risinger, J. I., Boyd, J., IonovY., , Perucho, M. and Kunkel, T. A. (1994) Defective mismatch repair in extracts of colorectal and endometrial cancer cell lines exhibiting microsatellite instability. *J. Biol. Chem.*, **269**, 14367-14370.

26. Larson, A. A., Kern, S., Sommers, R. L., Yokota, J., Cavenee, W. K. and Hampton, G. M. (1996) Analysis of replication error (RER+) phenotypes in cervical carcinoma. *Cancer Res.*, **56**, 1426-1431.

27. Johnson, E. S., Ma, P. C., Ota, I. M. and Varshavsky, A. (1995) A proteolytic pathway that recognizes ubiquitin as a degradation signal. *J Biol Chem.*, **270**, 17442-17456.

28. Band, V., Dalal, S., Delmolino, L. and Androphy, E. J. (1993) Enhanced degradation of p53 protein in HPV-6 and BPV-1 E6-immortalized human mammary epithelial cells. *EMBO J.*, **12**, 1847-1852.

29. Lima, C. D., Klein, M. G., Weinstien, I. B. and Hendrickson, W. A. (1996) Three-dimensional structure of human protein kinase C interacting protein 1, a member of the HIT family of proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 5357-5362.

30. Virgilio, L., Shuster, M., Gollin, S. M., Veronese, M. L., Ohta, M., Huebner, K. and Croce, C. M. (1996) FHIT gene alterations in head and neck squamous cell carcinomas. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 9770-9775.

31. Borodovsky, M. and McIninch, J. (1993) GenMark: Parallel gene recognition for both DNA strands. *Computers Chem.*, **17**, 123-133.

32. Milosavljevic, A. in *Imperial Cancer Research Fund Handbook of Genome Analysis* Vol. 1, Blackwell Scientific Publications, 1994).

33. Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Pizzuti, A., Richards, S., Victoria, M. F., Zhang, F. P. and et al., (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, **65**, 905-914.

34. Kremer, E. J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S. T., Sutherland, G. R. and Richards, R. I. (1991) Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Sci.*, **252**, 1711-1714.

35. Wagatsuma, M., Hashimoto, K. and Matsukura, T. (1990) Analysis of integrated human papillomavirus type 16 DNA in cervical cancers: amplification of viral sequences together with cellular flanking sequences. *J Virol.*, **64**, 813-821.

36. Wilke, C. M., Hall, B. K., Hoge, A., Paradee, W., Smith, D. I. and Glover, T. W. (1996) FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. *Hum. Mol. Genet.*, **5**, 187-195.

37. Cohen, S. and Lavi, S. (1996) Induction of circles of heterogeneous sizes in carcinogen-treated cells: two-dimensional gel analysis of circular DNA molecules. *Mol. Cell. Biol.*, **16**, 2002-2014.

38. Misra, R., Shih, A., Rush, M., Wong, E. and Schmid, C. W. (1987) Cloned extrachromosomal circular DNA copies of the human transposable element THE-1 are related predominantly to a single type of family member. *J. Mol. Biol.*, **196**, 233-243.

39. Motejlek, K., Schindler, D., Assum, G. and Krone, W. (1993) Increased amount and contour length distribution of small polydisperse circular DNA (spcDNA) in Fanconi anemia. *Mutat. Res.*, **293**, 205-214.

40. Sunnerhagen, P., Sjoberg, R. M. and Bjursell, G. (1989) Increase of extrachromosomal circular DNA in mouse 3T6 cells on perturbation of DNA synthesis: implications for gene amplification. *Somat. Cell. Mol. Genet.*, **15**, 61-70.

41. Paradee, W., Mullins, C., He, Z., Glover, T., Wilke, C., Opalka, B. and Schutte, J. (1995) Precise localization of aphidicolin-induced breakpoints on the short arm of human chromosome 3. *Genomics*, **27**, 358-361.

42. Klein, S., Zenvirth, D., Sherman, A., Ried, K., Rappold, G. and Simchen, G. (1996) Double-strand breaks on YACs during yeast meiosis may reflect meiotic recombination in the human genome. *Nat. Genet.*, **13**, 481-484.

43. Sozzi, G., Veronese, M. L., Negrini, M., Baffa, R., Cotticelli, M. G., Inoue, H., Pilotti, S., De Gregorio, L., Pastorino, U., Pierotti, M. A., OhtaM., , Huebner, K. and Croce, C. M. (1996) The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell*, **85**, 17-26.

44. Thiagalingam, S., Lisitsyn, N. A., Hamaguchi, M., Wigler, M. H., Willson, J. K. V., Markowitz, S. D., Leach, F. S., Kinzler, K. W. and Vogelstein, B. (1996) Evaluation of the FHIT gene in colorectal cancers. *Cancer Res.*, **56**, 2936-2939.

45. Yin, Y., Tainsky, M. A., Bischoff, F. Z., Strong, L. C. and Wahl, G. M. (1992) Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. *Cell*, **70**, 937-948.

46. Livingstone, L. R., White, A., Sprouse, J., Livanos, E., Jacks, T. and Tlsty, T. D. (1992) Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. *Cell*, **70**, 923-935.

47. Renault, B., Calistri, D., Buonsanti, G., Nanni, O., Amadori, D. and Ranzani, G. N. (1996) Microsatellite instability and mutations of p53 and TGF-beta RII genes in gastric cancer. *Hum. Genet.*, **In Press**.

48. Wu, C., Akiyama, Y., Imai, K., Miyake, S., Nagasaki, H., Oto, M., Okabe, S., Iwama, T., Mitamura, K. and Masumitsu, H. (1994) DNA alterations in cells from hereditary non-polyposis colorectal cancer patients. *Oncogene.*, **9**, 991-994.

49. Pintado, E., de Diego, Y., Hmadcha, A., Carrasco, M., Sierra, J. and Lucas, M. (1995) Instability of the CGG repeat at the FRAXA locus and variable phenotypic expression in a large fragile X pedigree. *J Med. Genet*, **32**, 907-908.

50. Nancarrow, J. K., Holman, K., Mangelsdorf, M., Hori, T., Denton, M., Sutherland, G. R. and Richards, R. I. (1995) Molecular basis of p(CCG)n repeat instability at the FRA16A fragile site locus. *Hum. Mol. Genet*, **4**, 367-372.

51. Bouffler, S., Silver, A., Papworth, D., Coates, J. and Cox, R. (1993) Murine radiation myeloid leukaemogenesis: relationship between interstitial telomere-like sequences and chromosome 2 fragile sites. *Genes. Chromosomes. Cancer.*, **6**, 98-106.

52. Usdin, K. and Woodford, K. J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Res.*, **23**, 4202-4209.

53. Mullokandov, M. R., Kholodilov, N. G., Atkin, N. B., Burk, R. D., Johnson, A. B. and Klinger, H. P. (1996) Genomic alterations in cervical carcinoma: losses of chromosome heterozygosity and human papilloma virus tumor status. *Cancer Res.*, **56**, 197-205.

54. Drabkin, H. A., Wright, M., Jonsen, M., Varkony, T., Jones, C., Sage, M., Gold, S. and Erickson, P. (1990) Development of a somatic cell hybrid mapping panel and molecular probes for human chromosome 3. *Genomics.*, **8**, 435-446.

55. Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A. and de Jong, P. J. (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.*, **6**, 84-89.

56. Grossberger, D. (1987) Minipreps of DNA from bacteriophage lambda. *Nucleic. Acids. Res.*, **15**, 6737.

57. Sambrook, J. Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A laboratory manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

58. Xu, Y., Einstein, J. R., Mural, R. J., Shah, M. & Uberbacher, E. C. in *Proceedings of the 2nd International Conference on Intelligent Systems for Molecular Biology* Vol. 1, 376-384 (AAAI Press, 1994).

METHODS

1. Nucleic acid sources and manipulations

Cell lines: Breast, colon and cervical carcinoma lines were obtained from the American Type Culture Collection. The cervical carcinoma cell line, CC19, was established as described (53). Lung tumor lines were obtained through the Colorado Lung Cancer SPORE Tissue Culture Core Laboratory. Normal cell lines included the human lymphoblastoid cell lines TL8229 and AG4103 and the E1A-immortalized human embryonic kidney cell line, E293. The somatic cell hybrids have been described previously (54).

Libraries: The gridded chromosome 3-specific cosmid library (55) was obtained from Lawrence Livermore National Laboratories. The 850A6 YAC subclone library was previously described (16). Fetal brain and adult kidney cDNA libraries were obtained from Clontech.

DNA and RNA isolations: DNA was isolated from cell lines by standard methods. Cosmid DNAs were isolated using alkaline lysis; preparations used for DNA sequence analysis were purified further by CsCl gradient centrifugation. DNA was isolated from single or pooled phage clones by the Grossberger method (56). RNA was isolated from cell lines when the cultures reached 90% of confluency using the RNA-STAT-60 kit from Tel-Test, Inc. (Friendswood, TX).

Library screening, contig assembly and hybridization analysis: The gridded cosmid library was spotted onto filters at high density (1536 clones per filter) and hybridized using standard techniques with a 370 kb MluI fragment derived from YAC 74B2. The 850A6 YAC phage library (16) was screened using the same probe. Resulting clones were assembled into a contiguous segment by hybridization analysis using end-probes and total inserts followed by analysis with the

software tool SEGMAP. Complete and partial digestion analyses (57) were used to restriction map the central 170 kb of the contig prior to DNA sequence analysis.

2. Replication error (RER) analysis: Biotinylated primers were obtained from Research Genetics, Huntsville, AL. PCR amplifications were performed with 40 ng template DNA utilizing hot start and touch down procedures. After separation on denaturing polyacrylamide gels, PCR products were detected using the New England BioLabs Phototope™ Detection Kit. Alleles were scored by visual inspection of band patterns.

3. RT-PCR Analysis of FHIT gene expression RT-PCR was performed using primers FHIT-5' (5'-CTCGAATTCTTAGACCCTCTATAAAAGC-3') and FHIT-3' (5'CTGATTCAGTTCCTCTTG-3') derived from non-coding exons 4 and 10, respectively. First strand synthesis was accomplished with 1-3 μg total RNA and the Superscript II kit (Life Technologies). Subsequent amplification utilized one-fifth of the reverse transcriptase reaction together with the FHIT primers. Standard PCR conditions of 94°C denaturation (1 min), 55°C annealing (1 min) and 72°C elongation for 35 cycles were employed. PCR products were subcloned into the EcoRI site of pBlusecript II SK+ using an introduced EcoRI site present in the FHIT-5' primer and a natural site located 21 bp downstream of the FHIT stop codon.

4. DNA sequence determination. Cosmid and phage clones were sequenced using a random shotgun subcloning and end-sequencing strategy. Clone DNA was sonicated and size selected by LMP-agarose gel electrophoresis. Recovered fragments of 1-2 kb were end-repaired with Klenow fragment of E. coli DNA polymerase I and T4 DNA polymerase, ligated into the phosphatased EcoRV site of pBluescript II and transformed into E. coli DH10B. Amp^R /β-Gal^- subclones were grown in 3 mls TB for isolation of sequencing templates. To eliminate vector sequences from the subclone library, phage inserts were amplified by long range PCR. Inserts were gel purified, ^{32}-P

labeled and hybridized to subclone libraries. Positive subclones were sequenced using an ABI 373 or 377 and the ABI Prism dye terminator cycle sequencing kit. Chromatograms were transferred to a SUN workstation, analyzed and assembled using PHRED and PHRAP (from Dr. Phil Green). Gaps were closed by primer walking. Based on independently obtained overlapping contigs, as well as analysis of cosmid vector sequences in some subclones, sequencing accuracy was at least 99.6%.

5. Sequence analysis. Assembled sequences were analyzed for database similarities by BLASTN and BLASTX, searching the nr and dbest databases with default parameters. Strongly similar sequences ($p < e^{-30}$) were retrieved from GenBank for further analysis. Homologies to repeat sequences were found using Pythia (32). Potential exons were identified using GRAIL2 (58) and GeneMark (31) programs. GeneMark was run using 5th order matrices trained on sequences with GC content similar to the sequences checked. Sequence alignments were prepared using the FASTA program from the Wisconsin Package (GCG).

Figure Legends

Figure 1.

A. Schematic showing position of the homozygous deletion region with respect to YACs 850A6, 65E7 and 74B2. Positions of MluI and selected XhoI restriction sites are indicated. XhoI sites were not determined for YAC 850A6.

B. Cosmid and lambda clone contig spanning the homozygous deletion region within 3p14. Solid horizontal lines represent cosmid (c) and lambda (λ) inserts, as indicated, along with cleavage sites for EcoRI (short vertical bars), XhoI (X) and SalI (S) which were mapped within the central 170 kb. The MluI site corresponding to the site in YAC 74B2 is present in c55D12 and c84F12, although we have not determined its precise location. Positions of selected markers,

breakpoints and integration sites for HPV-16 and pSV2neo are indicated across the top. Parts B and C of this figure are drawn to the same scale and positions correspond exactly between the two.

C. Homozygous deletions in tumors and normal genomic DNAs. Horizontal lines denote the extent of deletion in the indicated cell lines. Parallel lines at the ends of each deletion indicate where precise boundaries were not determined. HeLa and MDA231 contain discontinuous deletions indicated by interrupted lines. Exon 5 and the direction of transcription for the FHIT gene are indicated at the bottom.

Figure 2.

A. Southern analysis of cervical carcinoma cell lines. Lanes 1 through 7 contain DNA from seven cervical lines, as indicated; lane 8 is a normal human DNA control. DNA samples were digested with EcoRI and hybridized with c136C3. Three cell lines were homozygously deleted for c136C3 while the arrow indicates a rearranged band present in CC19.

B. Southern analysis of colon carcinoma cell lines. Control lanes (1-4) contain respectively, λ HindIII size marker, HindIII digested human DNA, the chromosome 3-specific hybrids 2A3CT and UCH12. Hybrid 2A3CT contains a single copy of chromosome 3 deleted for all sequences distal to 3p21.3. The chromosome retained in UCH12 is deleted for the entire short (p) arm. Lanes 5 through 18 contain HindIII digested DNAs from 12 colon carcinoma cell lines (note, DLD-1 and HCT-15 are identical). The Southern blot was hybridized with c61G12 which contains D3S1300 and FHIT exon 5. Homozygous deletions were observed in three cell lines; altered bands, which may be indicative of rearrangements, were present in 4 additional lines.

Figure 3. Deletions in genomic DNAs from normal sources. DNA from several hybrids and YAC clones were analyzed with probes from the homozygous deletion region. In panel A, DNA from 7 chromosome 3 hybrids and a deletion variant of YAC clone 74B2 were digested with HindIII and hybridized with c61G12. The variant band in hybrid H3-4 may represent a polymorphism. The seven hybrids all retain 3p14.2 by cytogenetic and molecular genetic analyses while missing other

specific regions of chromosome 3. Panel B; DNA samples from YACs 74B2 and the deletion variant 74B2Δ were digested with EcoRI and hybridized with c31E1. Four homologous bands present in YAC 74B2 were missing in 74B2Δ (arrows). In panel C, EcoRI digested DNA from the hybrid A5-4 (lane 4) was compared to human (lane 1) and hybrid 3;8/4-1 (lane 3) using c31E1.

Figure 4. Panel A. RT-PCR analysis of FHIT expression in tumor cell lines. Primers designed to amplify the coding region of FHIT were used to amplify cDNAs prepared from 5 cervical carcinoma lines, one renal carcinoma line (KRC/Y) and an E1A transformed embryonal kidney line (E293), as indicated. Normal products of 638 bp (arrow) were observed in three cell lines, including CC19 with a 40 kb homozygous deletion. A normal sized product was also observed in MS751, although it is amplified weakly. In panel B, evidence is presented for alternative splicing of FHIT mRNA in normal tissues. RT-PCR analysis of FHIT expression in fetal brain, adult kidney, E293 cells (whole cells and cytoplasmic fraction), lung tumor A549 and kidney carcinoma cell line KRC/Y were performed as in panel A. Both normal sized products (arrow marked N) and smaller products (left arrow) were observed.

Figure 5. Assembled sequence data totaling 110,435 basepairs. GeneMark and GRAIL2 predicted exons are indicated (filled boxes = direct strand; open boxes = complementary strand). Repetitive sequences homologous to L1 and Alu elements are designated by boxes filled with diagonal lines and shading, respectively. Other repetitive elements are labeled. Unique features of the sequence are indicated on the lowest row.

Figure 6. FASTA alignment of 585 bp of spcDNA clone (X96885) and FRA3B.

Table I. Results of microsatellite analysis and 3p14 deletions in selected cell lines.

UNPUBLISHED DATA

Table II - Summary of Sequence Features. The sequences used for comparison were derived from Xq27.3 (contains the fragile X region, GenBank-L29074) and 4p16.3 (in the region of Huntington's disease, GenBank-Z69837). The sequence from 3p21.3 was downloaded from ftp://genome.wustl.edu/pub/gsc1/sequence/st.louis/human/shotgun/3/H_LUCA14.seq.

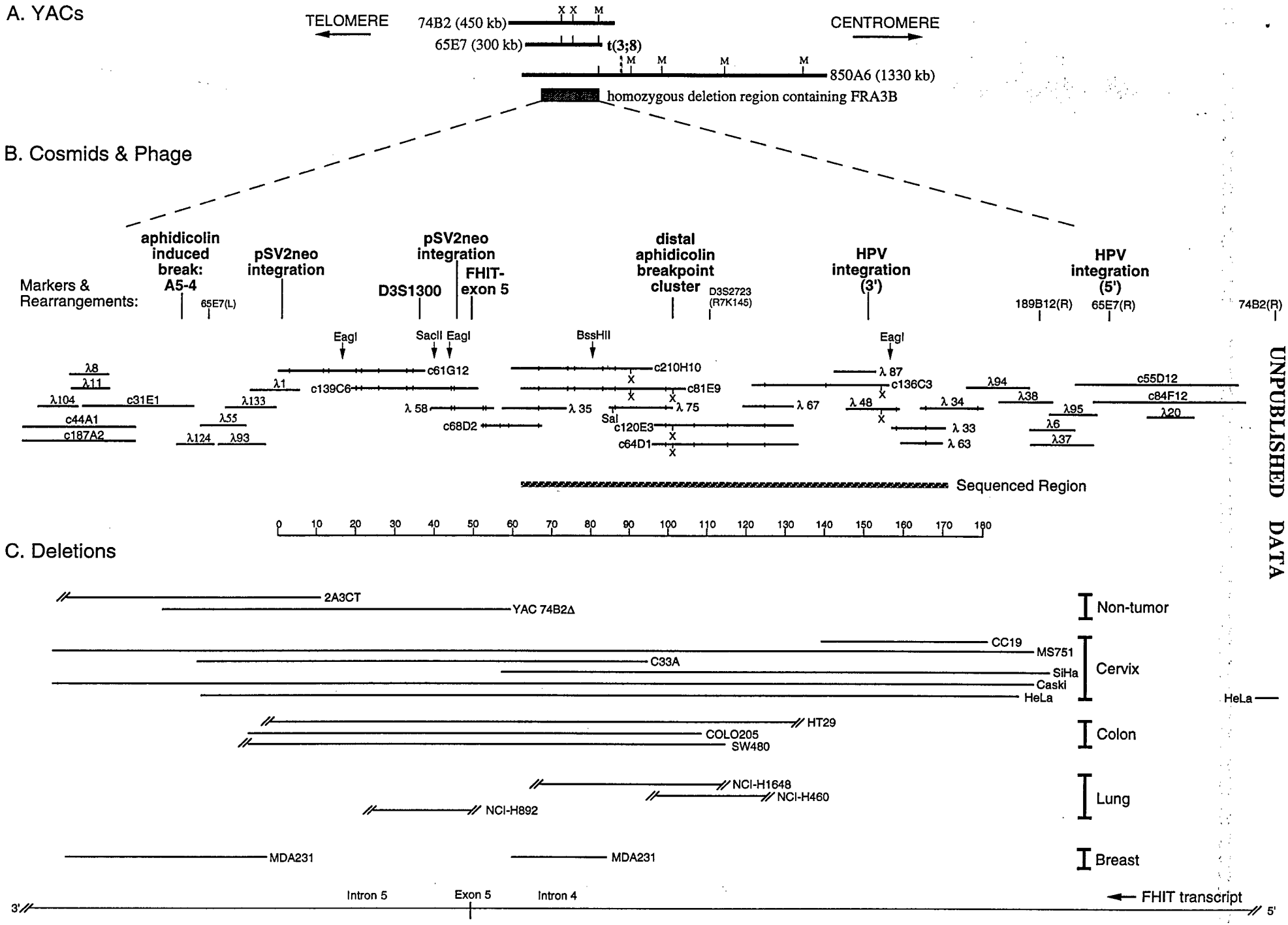


FIGURE 2

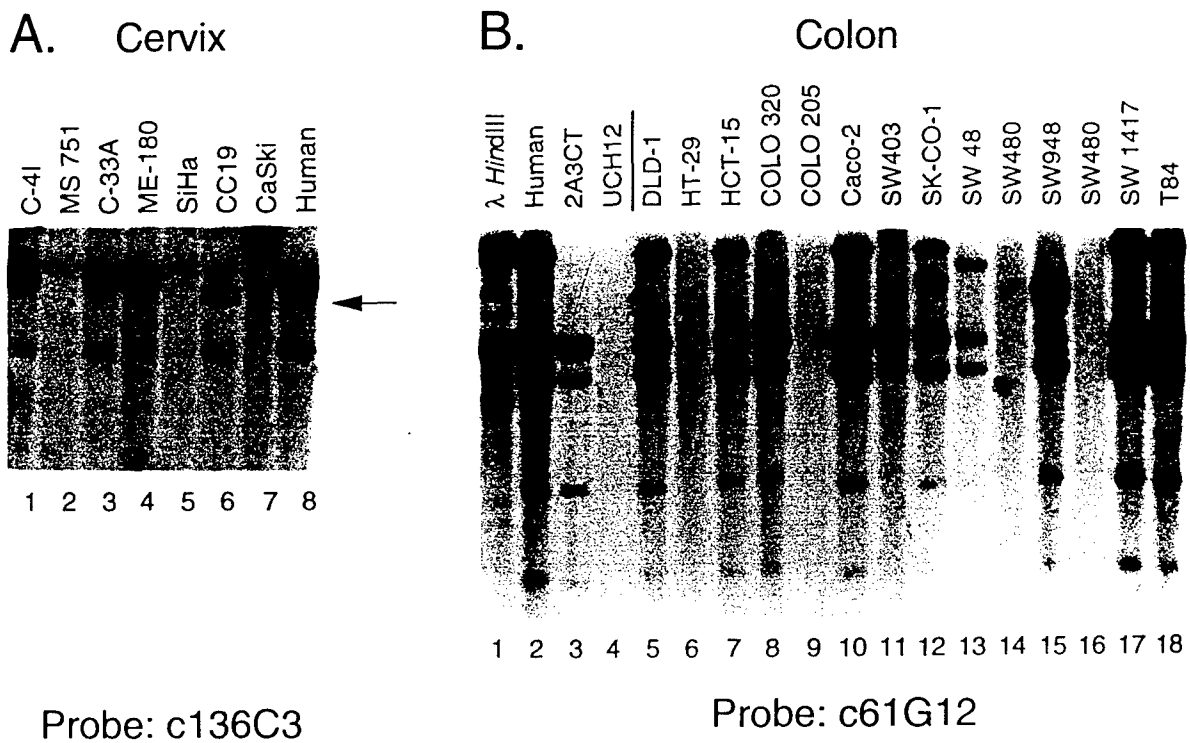
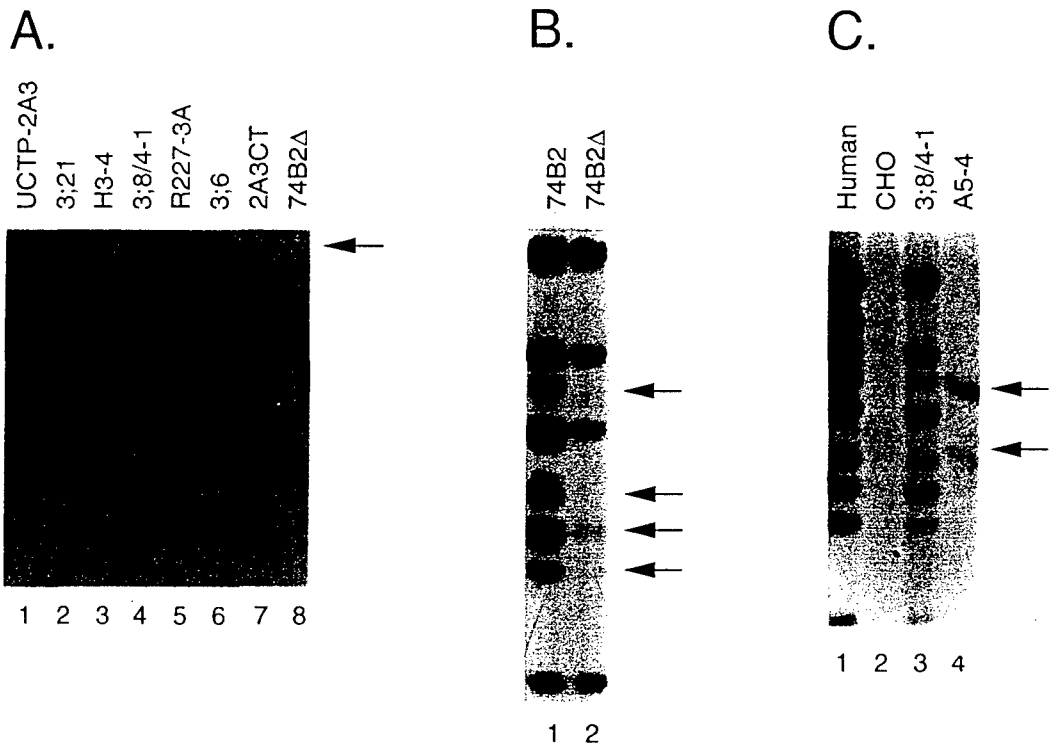

A.
Cervix
C-4I
MS 751
C-33A
ME-180
SiHa
CC19
CaSki
Human
1 2 3 4 5 6 7 8
Probe: c136C3
B.
Colon
λ HindIII
Human
2A3CT
UCH12
DLD-1
HT-29
HCT-15
COLO 320
COLO 205
Caco-2
SW403
SK-CO-1
SW 48
SW480
SW948
SW480
SW 1417
T84
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Probe: c61G12

FIGURE 3



Probe: c61G12 Probe: c31E1 Probe: c31E1

FIGURE 4A

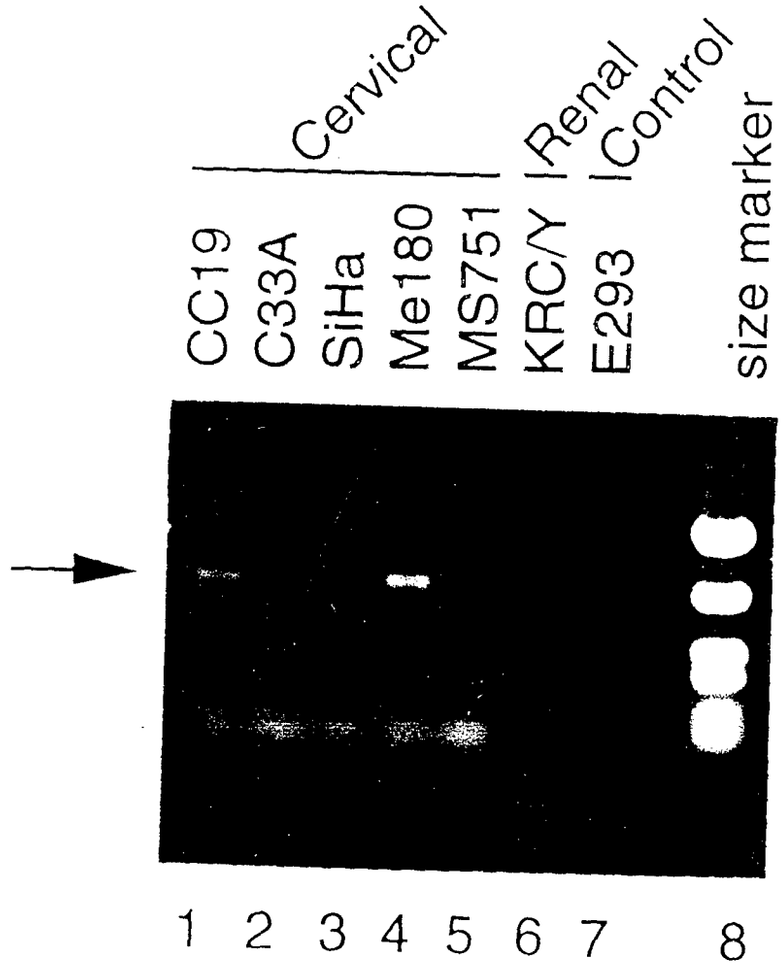
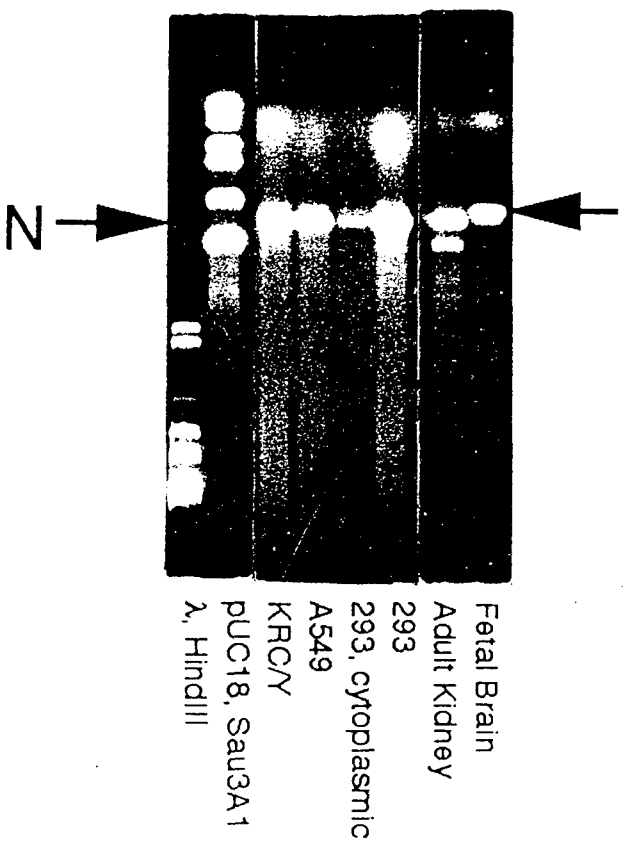
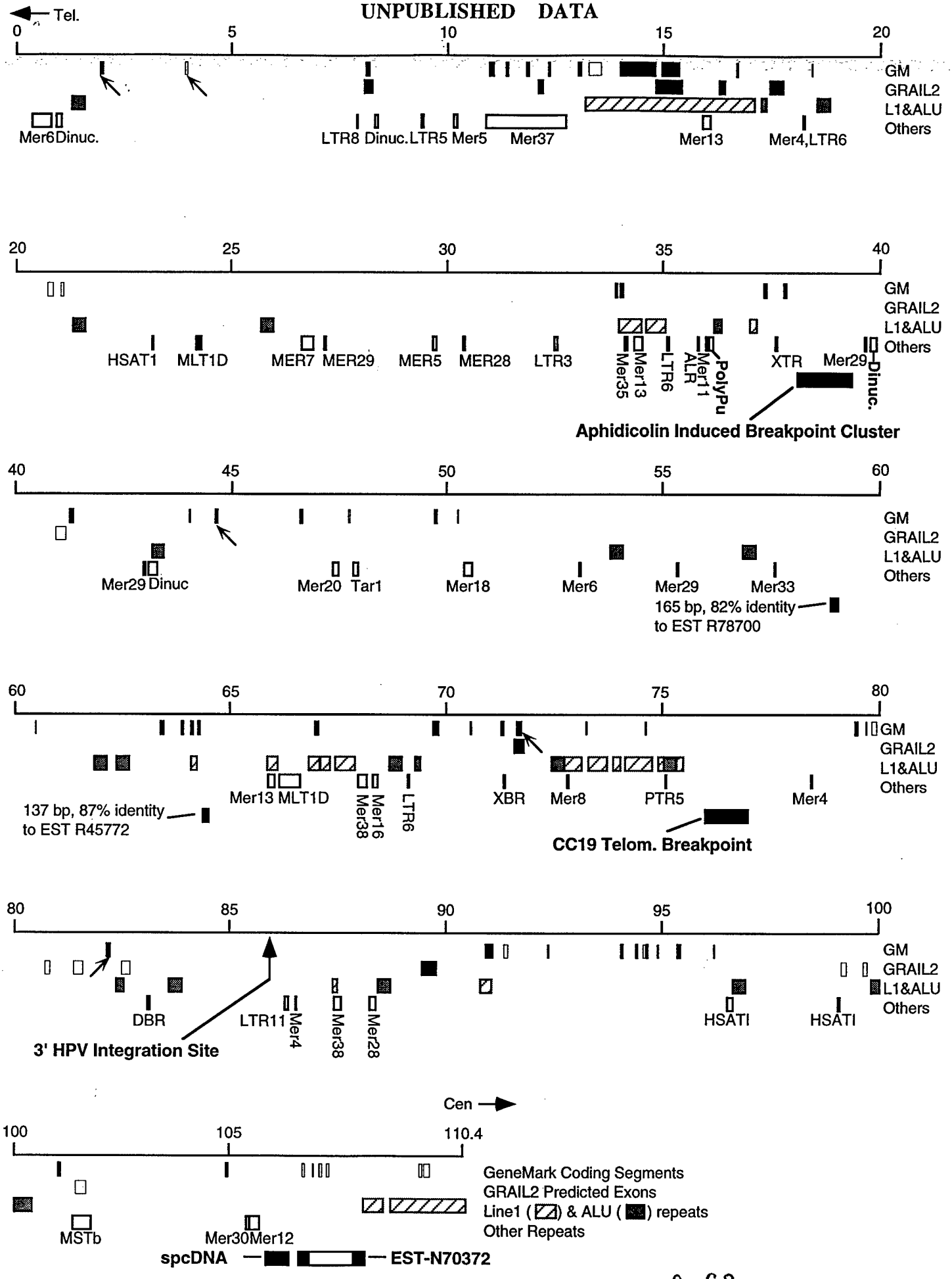

Cervical
Renal
Control
CC19
C33A
SiHa
Me180
MS751
KRC/Y
E293
size marker
1 2 3 4 5 6 7 8

FIGURE 4B


N
λ, HindIII
pUC18, Sau3A1
KRC/Y
A549
293, cytoplasmic
293
Adult Kidney
Fetal Brain


Tel.
0
5
10
15
20
GM
GRAIL2
L1&ALU
Others
Mer6Dinuc.
LTR8 Dinuc.LTR5 Mer5
Mer37
Mer13
Mer4,LTR6
20
25
30
35
40
HSAT1
MLT1D
MER7 MER29
MER5
MER28
LTR3
Mer35
Mer13
LTR6
Mer11 ALR
PolyPu
XTR
Mer29
Dinuc.
Aphidicolin Induced Breakpoint Cluster
40
45
50
55
60
Mer29 Dinuc
Mer20 Tar1
Mer18
Mer6
Mer29
Mer33
165 bp, 82% identity to EST R78700
60
65
70
75
80
Mer13 MLT1D
Mer38
Mer16
LTR6
XBR
Mer8
PTR5
Mer4
137 bp, 87% identity to EST R45772
CC19 Telom. Breakpoint
80
85
90
95
100
DBR
LTR11
Mer4
Mer38
Mer28
HSATI
HSATI
3' HPV Integration Site
Cen
100
105
110.4
MSTb
Mer30Mer12
spcDNA
EST-N70372
GeneMark Coding Segments
GRAIL2 Predicted Exons
Line1 ( ) & ALU ( ) repeats
Other Repeats

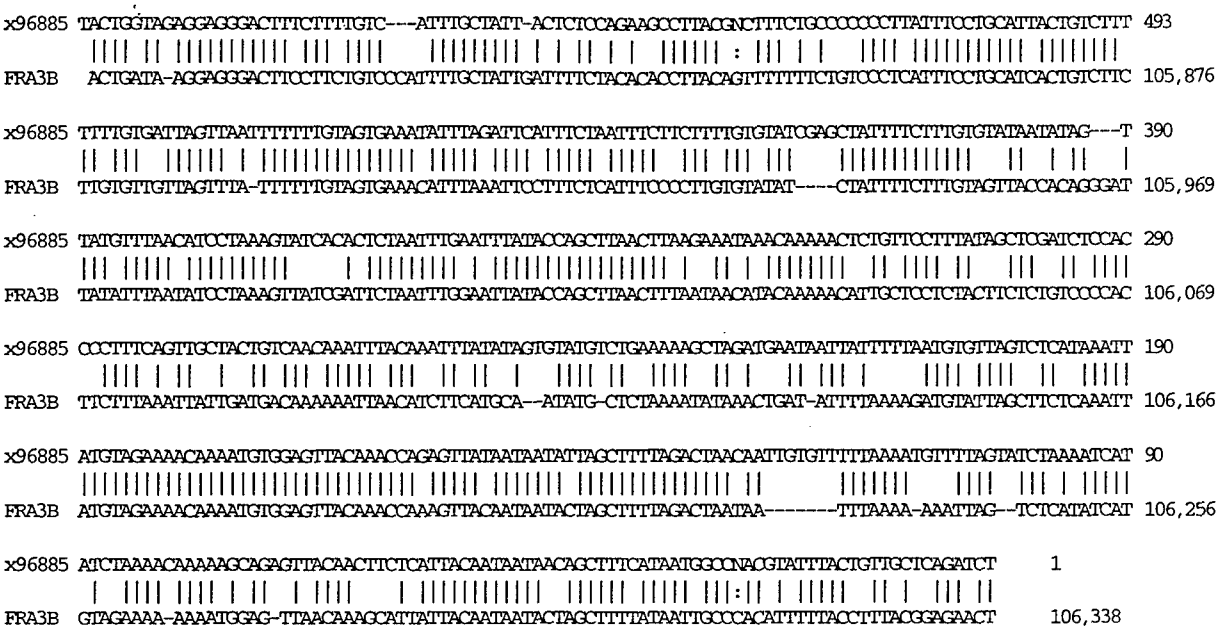

x96885 TACTGGTAGAGGAGGGACTTTCTTTTGTC---ATTTGCTATT-ACTCTCCAGAAGCCTTACGNCTTTCTGCCCCCCCCTTATTTCCTGCATTACTGTCTTT 493
FRA3B ACTGATA-AGGAGGGACTTCCTTCTGTCCCATTTTGCTATTGATTTTCTACACACCTTACAGTTTTTTTCTGTCCCTCATTTCCTGCATCACTGTCTTC 105,876
x96885 TTTTGTGATTAGTTAATTTTTTTGTAGTGAAATATTTAGATTCATTTCTAATTTCTTCTTTTGTGTATCGAGCTATTTTCTTTGTGTATAATATAG---T 390
FRA3B TTGTGTTGTTAGTTTA-TTTTTTGTAGTGAAACATTTAAATTCCTTTCTCATTTCCCCTTGTGTATAT----CTATTTTCTTTGTAGTTACCACAGGGAT 105,969
x96885 TATGTTTAACATCCTAAAGTATCACACTCTAATTTGAATTTATACCAGCTTAACTTAAGAAATAAACAAAAACTCTGTTCCTTTATAGCTCGATCTCCAC 290
FRA3B TATATTTAATATCCTAAAGTTATCGATTCTAATTTGGAATTATACCAGCTTAACTTTAATAACATACAAAAACATTGCTCCTCTACTTCTCTGTCCCCAC 106,069
x96885 CCCTTTCAGTTGCTACTGTCAACAAATTTACAAATTTATATAGTGTATGTCTGAAAAAGCTAGATGAATAATTATTTTTAATGTGTTAGTCTCATAAATT 190
FRA3B TTCTTTAAATTATTGATGACAAAAAATTAACATCTTCATGCA--ATATG-CTCTAAAATATAAACTGAT-ATTTTAAAAGATGTATTAGCTTCTCAAATT 106,166
x96885 ATGTAGAAAACAAAATGTGGAGTTACAAACCAGAGTTATAATAATATTAGCTTTTAGACTAACAATTGTGTTTTTAAAATGTTTTAGTATCTAAAATCAT 90
FRA3B ATGTAGAAAACAAAATGTGGAGTTACAAACCAAAGTTACAATAATACTAGCTTTTAGACTAATAA-------TTTAAAA-AAATTAG--TCTCATATCAT 106,256
x96885 ATCTAAAACAAAAAGCAGAGTTACAACTTCTCATTACAATAATAACAGCTTTCATAATGGCCNACGTATTTACTGTTGCTCAGATCT 1
FRA3B GTAGAAAA-AAAATGGAG-TTAACAAAGCATTATTACAATAATACTAGCTTTTATAATTGCCCACATTTTTACCTTTACGGAGAACT 106,338

Table I.

Cell Line	Mismatch Repair Deficiency / Microsatellite Instability *	3p14 status	Reference
Colorectal Carcinoma †			
HT-29	intact	deletion	this reference
SW-480	intact	deletion	Umar et al., 1994; this reference
SW-403	intact	deletion	this reference
SW-948	intact	rearranged	this reference
SW-48	defective	no deletion	hMLH1 - Boyer et al., 1995; this reference
DLD1/HCT15	defective	no deletion	Pol δ {da Costa et al., 1995} GTBP reviewed in {Bhattacharyya et al., 1995}
Cervical Carcinoma ††			
C-41	intact	deletion	Larson et al., 1996 **
Caski	intact	deletion	Larson et al., 1996 **
SIHA	intact	deletion	Larson et al., 1996 **
HeLa	intact	deletion	Boyer et al., 1995; Umar et al., 1994
CC19	NT	deletion	
MS-751	intact	deletion	Larson et al., 1996 **
C-33A	defective	deletion	Larson et al., 1996 **, this reference
ME-180	intact	no deletion	Larson et al., 1996 **

* Defective refers to a reported deficiency in mismatch repair or microsatellite instability.
† Overall, deletions (5) or rearrangements (1) were identified in 6/12 colorectal carcinoma DNAs. In addition to those listed, deletions occurred in COLO-205 and COLO-320. No deletions were identified in SKCO-1, CaCo-2, SW-1417 and T-84.
†† represents all cervical carcinoma samples examined
NT = not tested
** Larson et al., 1996 and personal communication from Dr. Garret M. Hampton

	Sequence Length	GC Content	Total Repeats	Alu	Line	Mer	Predicted Exons	Confirmed Genes
FRA3B (110.4 kb Sequence)	110,434	38.9%	20.2%	4.9%	9.0%	3.9%	3.4%	0.0%
FRA3B (λ33 only)	13,121	37.0%	26.6%	3.5%	17.5%	2.1%	2.3%	0.0%
Xq27.3(Fragile X)	152,351	39.0%	26.3%	5.9%	16.3%	3.0%	3.4%	2.8%
3p21.3(HLuca_14)	36,597	58.0%	19.4%	18.1%	0.0%	0.4%	9.5%	12.9%
4p16.3(Huntington's)	32,100	54.9%	16.1%	10.8%	3.5%	1.4%	5.6%	7.6%

DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH AND MATERIEL COMMAND
504 SCOTT STREET
FORT DETRICK, MARYLAND 21702-5012

REPLY TO
ATTENTION OF:

MCMR-RMI-S (70-1y)

1 JUN 2001

MEMORANDUM FOR Administrator, Defense Technical Information Center (DTIC-OCA), 8725 John J. Kingman Road, Fort Belvoir, VA 22060-6218

SUBJECT: Request Change in Distribution Statement

1. The U.S. Army Medical Research and Materiel Command has reexamined the need for the limitation assigned to technical reports. Request the limited distribution statement for reports on the enclosed list be changed to "Approved for public release; distribution unlimited." These reports should be released to the National Technical Information Service.

2. Point of contact for this request is Ms. Judy Pawlus at DSN 343-7322 or by e-mail at judy.pawlus@det.amedd.army.mil.

FOR THE COMMANDER:

Encl

PHYLLIS M. RINEHART
Deputy Chief of Staff for
Information Management

DAMD17-94-J-4413	ADB261602
DAMD17-96-1-6112	ADB233138
DAMD17-96-1-6112	ADB241664
DAMD17-96-1-6112	ADB259038
DAMD17-97-1-7084	ADB238008
DAMD17-97-1-7084	ADB251635
DAMD17-97-1-7084	ADB258430
DAMD17-98-1-8069	ADB259879
DAMD17-98-1-8069	ADB259953
DAMD17-97-C-7066	ADB242427
DAMD17-97-C-7066	ADB260252
DAMD17-97-1-7165	ADB249668
DAMD17-97-1-7165	ADB258879
DAMD17-97-1-7153	ADB248345
DAMD17-97-1-7153	ADB258834
DAMD17-96-1-6102	ADB240188
DAMD17-96-1-6102	ADB257406
DAMD17-97-1-7080	ADB240660
DAMD17-97-1-7080	ADB252910
DAMD17-96-1-6295	ADB249407
DAMD17-96-1-6295	ADB259330
DAMD17-96-1-6284	ADB240578
DAMD17-96-1-6284	ADB259036
DAMD17-97-1-7140	ADB251634
DAMD17-97-1-7140	ADB259959
DAMD17-96-1-6066	ADB235510
DAMD17-96-1-6029	ADB259877
DAMD17-96-1-6020	ADB244256
DAMD17-96-1-6023	ADB231769
DAMD17-94-J-4475	ADB258846
DAMD17-99-1-9048	ADB258562
DAMD17-99-1-9035	ADB261532
DAMD17-98-C-8029	ADB261408
DAMD17-97-1-7299	ADB258750
DAMD17-97-1-7060	ADB257715
DAMD17-97-1-7009	ADB252283
DAMD17-96-1-6152	ADB228766
DAMD17-96-1-6146	ADB253635
DAMD17-96-1-6098	ADB239338
DAMD17-94-J-4370	ADB235501
DAMD17-94-J-4360	ADB220023
DAMD17-94-J-4317	ADB222726
DAMD17-94-J-4055	ADB220035
DAMD17-94-J-4112	ADB222127
DAMD17-94-J-4391	ADB219964
DAMD17-94-J-4391	ADB233754